



(12) 发明专利申请

(10) 申请公布号 CN 118451194 A

(43) 申请公布日 2024. 08. 06

(21) 申请号 202280069737.1

(22) 申请日 2022.10.17

(30) 优先权数据

PCT/EP2021/078828 2021.10.18 EP

(85) PCT国际申请进入国家阶段日

2024.04.16

(86) PCT国际申请的申请数据

PCT/EP2022/078869 2022.10.17

(87) PCT国际申请的公布数据

W02023/066874 EN 2023.04.27

(71) 申请人 生物技术公司

地址 德国

申请人 约翰·古腾堡大学美因兹医学大学
转化肿瘤学公司

(72) 发明人 蒂姆·拜塞特

马里奥·佩尔科维奇

斯特凡尼·加夫莱塔

帕特里克·索恩 西尔克·布里尔

埃费林·内特

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

专利代理师 张福誉 韩晓帆

(51) Int.Cl.

C12Q 1/6811 (2006.01)

C12Q 1/6876 (2006.01)

权利要求书4页 说明书84页

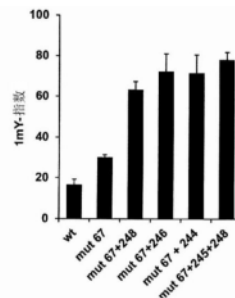
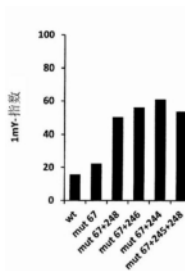
序列表(电子公布) 附图24页

(54) 发明名称

用于确定用于提高经修饰可复制RNA功能的突变的方法以及相关组合物及其用途

(57) 摘要

本发明涉及用于恢复或提高含有经修饰核苷酸的可复制RNA的复制和/或翻译能力的方法。所述方法包括在所述可复制RNA中鉴定核苷酸变化,所述核苷酸变化补偿了含有经修饰核苷酸的可复制RNA分子的复制和/或翻译能力的降低。本发明还涉及并入这样鉴定出的核苷酸变化的含有经修饰核苷酸的可复制RNA分子以及这样的可复制RNA分子在治疗中的用途。



1. 用于在包含经修饰核苷酸的可复制RNA (rRNA) (经修饰rRNA) 中鉴定序列变化的方法, 所述经修饰核苷酸至少部分恢复或提高经修饰rRNA的功能, 所述方法包括:

(a) 用编码目的基因的经修饰rRNA转染细胞, 所述细胞表达RNA依赖性RNA聚合酶(复制酶); 以及

(b) 鉴定表达所述目的基因的经转染细胞中存在的编码所述目的基因的rRNA分子内的序列变化。

2. 根据权利要求1所述的方法, 其中所述序列变化是相对于步骤(a) 中转染的经修饰rRNA的序列来确定的。

3. 根据权利要求1或2所述的方法, 其中步骤(a) 还包括在转染之后, 从表达所述目的基因的经转染细胞中分离编码所述目的基因的rRNA分子以提供分离的rRNA分子。

4. 根据权利要求3所述的方法, 其中步骤(a) 还包括用编码所述目的基因并包含所述经修饰核苷酸的rRNA分子的后代对表达所述复制酶的细胞进行转染的步骤, 所述后代是通过复制所述分离的rRNA分子来提供的。

5. 根据权利要求4所述的方法, 所述方法还包括重复权利要求3和4所述的步骤。

6. 根据权利要求5所述的方法, 其中所述步骤重复n次。

7. 根据权利要求6所述的方法, 其中n是至少为1的整数。

8. 根据权利要求4至7中任一项所述的方法, 其中复制所述分离的rRNA分子包括(i) 对编码所述目的基因的所述分离的rRNA分子进行逆转录, 以形成能够被体外转录的DNA分子; 以及(ii) 在相同的经修饰核苷酸的存在下体外转录所述DNA分子, 以产生包含所述经修饰核苷酸并编码所述目的基因的经修饰rRNA分子的后代。

9. 根据权利要求1至8中任一项所述的方法, 其还包括以下步骤:

(c) 将至少一个鉴定出的序列变化并入到步骤(a) 的经修饰rRNA中;

(d) 用步骤(c) 中产生的经修饰rRNA对表达所述复制酶的细胞进行转染; 以及

(e) 鉴定从步骤(d) 中产生的表达所述目的基因的经转染细胞中分离的rRNA分子内的序列变化。

10. 根据权利要求1至9中任一项所述的方法, 其中所述经修饰rRNA中经修饰核苷酸的数目是所述rRNA中核苷酸总数目的至少10%、20%、30%、40%、50%、60%、70%、80%、90%或100%。

11. 根据权利要求10所述的方法, 其中所述经修饰rRNA中经修饰核苷酸的数目是所述rRNA中核苷酸总数目的至少30%。

12. 根据权利要求10所述的方法, 其中所述经修饰rRNA中经修饰核苷酸的数目是所述rRNA中核苷酸总数目的至少50%。

13. 根据权利要求1至12中任一项所述的方法, 其中所述经修饰核苷酸是经修饰的尿苷残基、经修饰的腺嘌呤残基、经修饰的鸟嘌呤残基、经修饰的胞嘧啶残基, 或者前述两种或更多种的任何组合。

14. 根据权利要求1至13中任一项所述的方法, 其中所述经修饰核苷酸是经修饰的尿苷。

15. 根据权利要求14所述的方法, 其中所述经修饰的尿苷是假尿苷。

16. 根据权利要求15所述的方法, 其中所述假尿苷是N1-甲基-假尿苷。

17. 根据权利要求16所述的方法, 其中在体外转录的经修饰rRNA中N1-甲基-假尿苷残基的数目是尿苷残基总数目的至少10%、20%、30%、40%、50%、60%、70%、80%、90%或100%。

18. 根据权利要求17所述的方法, 其中在所述体外转录的经修饰rRNA中N1-甲基-假尿苷残基的数目是尿苷残基总数目的至少30%。

19. 根据权利要求17所述的方法, 其中在所述体外转录的经修饰rRNA中N1-甲基-假尿苷残基的数目是尿苷残基总数目的至少50%。

20. 根据权利要求1至19中任一项所述的方法, 其中所述复制酶来源于自复制RNA病毒, 优选甲病毒。

21. 根据权利要求20所述的方法, 其中所述甲病毒是SFV或VEEV或EEEV或辛德比斯病毒。

22. 根据权利要求1至21中任一项所述的方法, 其中所述rRNA包含不具有起始密码子的5' 调节区。

23. 根据权利要求1至22中任一项所述的方法, 其中所述复制酶和所述rRNA中复制酶复制所需的调节序列来源于相同的甲病毒或来源于不同的甲病毒。

24. 根据权利要求1至23中任一项所述的方法, 其中编码的目的基因是细胞表面表达的蛋白质或荧光素酶或GFP。

25. 根据权利要求1至24中任一项所述的方法, 其中所述复制酶在细胞中的表达是瞬时表达或组成型表达。

26. 根据权利要求1至25中任一项所述的方法, 其中在所述rRNA的转染之前、同时或之后用编码所述复制酶的核酸分子转染所述细胞, 优选地其中所述核酸分子是RNA。

27. 根据权利要求1至26中任一项所述的方法, 其中步骤(a) 包括用编码目的基因的自扩增RNA转染细胞, 该自复制RNA包含至少一个经修饰核苷酸。

28. 根据权利要求1至27中任一项所述的方法, 其中部分恢复或提高的所述rRNA的功能是由所述复制酶复制的能力和/或翻译的能力。

29. 根据权利要求1至28中任一项所述的方法, 其中所述功能恢复了至少10%或提高了至少10%。

30. 根据权利要求1至29中任一项所述的方法, 其中鉴定所述序列变化的步骤包括对所述rRNA分子的至少一部分进行测序。

31. 根据权利要求30所述的方法, 其中所测序的rRNA部分是5' 调节区。

32. 根据权利要求30所述的方法, 其中所测序的rRNA部分是3' 调节区。

33. 根据权利要求30至32中任一项所述的方法, 其中所述序列变化是通过分析由测序产生的序列特异性读数频率来确定的。

34. 根据权利要求1至33中任一项所述的方法, 其中进行所述方法以提高包含至少一个经修饰核苷酸的自扩增RNA的复制效率。

35. 根据权利要求1至34中任一项所述的方法, 其中所述方法还包括通过并入至少一个鉴定出的序列变化来修饰包含相同经修饰核苷酸的rRNA或自复制RNA的核苷酸序列, 所述鉴定出的序列变化部分恢复或提高所述包含经修饰核苷酸的rRNA或自复制RNA的功能。

36. 可复制RNA (rRNA) 分子, 其包含至少一个经修饰核苷酸并且通过根据权利要求1至

35中任一项所述的方法获得。

37.可复制RNA (rRNA),其具有至少部分恢复或提高经修饰rRNA功能的序列变化,所述序列变化通过包括以下的方法获得:

(a)用编码目的基因的经修饰rRNA转染细胞,所述细胞表达RNA依赖性RNA聚合酶(复制酶);以及

(b)鉴定表达所述目的基因的经转染细胞中存在的编码所述目的基因的rRNA分子内的序列变化。

38.可复制RNA (rRNA) 分子,其包含自复制RNA病毒的经修饰5' 调节区,所述经修饰调节区包含位于所述5' 调节区 (SEQ ID NO:2) 的第67、244、245、246、248位中的一个或多个处的点突变。

39.根据权利要求38所述的rRNA分子,其中所述自复制RNA病毒是甲病毒。

40.根据权利要求38或39所述的rRNA分子,其中所述rRNA分子包含位于所述5' 调节区 (SEQ ID NO:2) 的第67位处以及第244、245、246、248位中的一个或多个处的点突变。

41.根据权利要求38至40中任一项所述的rRNA分子,其中所述5' 调节区还包含位于所述5' 调节区 (SEQ ID NO:2) 的第4位处的点突变。

42.根据权利要求38至41中任一项所述的rRNA分子,其中所述点突变位于第67位处。

43.根据权利要求38至41中任一项所述的rRNA分子,其中所述点突变位于第67和244位处。

44.根据权利要求38至41中任一项所述的rRNA分子,其中所述点突变位于第67和246位处。

45.根据权利要求38至41中任一项所述的rRNA分子,其中所述点突变位于第67和248位处。

46.根据权利要求38至41中任一项所述的rRNA分子,其中所述点突变位于第67、245和248位处。

47.根据权利要求38至41中任一项所述的rRNA分子,其中所述点突变位于第4和67位处。

48.根据权利要求38至41中任一项所述的rRNA分子,其中所述点突变位于第4、67和244位处。

49.根据权利要求38至41中任一项所述的rRNA分子,其中所述点突变位于第4、67和246位处。

50.根据权利要求38至41中任一项所述的rRNA分子,其中所述点突变位于第4、67和248位处。

51.根据权利要求38至41中任一项所述的rRNA分子,其中所述点突变位于第4、67、245和248位处。

52.根据权利要求38至51中任一项所述的rRNA分子,其中所述点突变是G4A、A67C、G244A、C245A、G246A或C248A。

53.根据权利要求36至52中任一项所述的rRNA分子,其中所述自复制RNA病毒是SFV或EEEV或VEEV或辛德比斯病毒。

54.根据权利要求36至53中任一项所述的rRNA分子,其还包含一个或多个编码区。

55. 根据权利要求36至54中任一项所述的rRNA分子,其编码目的基因。
56. 根据权利要求55所述的rRNA分子,其中所编码的目的基因是抗原或治疗性蛋白质或核酸,优选肿瘤、病毒、细菌或真菌抗原或者变应原。
57. 根据权利要求36至56中任一项所述的rRNA分子,其编码RNA依赖性RNA聚合酶(复制酶),优选甲病毒复制酶。
58. 根据权利要求36至57中任一项所述的rRNA分子,其包含5' 调节区并且其编码复制酶,其中所述5' 调节区和编码的复制酶来源于相同的自复制RNA病毒,优选相同的甲病毒。
59. 根据权利要求36至57中任一项所述的rRNA分子,其包含5' 调节区并且其编码复制酶,其中所述5' 调节区和编码的复制酶来源于不同的自复制RNA病毒。
60. 根据权利要求36至59中任一项所述的rRNA分子,其是自扩增RNA。
61. 根据权利要求36至60中任一项所述的rRNA分子,其是体外转录的rRNA。
62. 根据权利要求36至61中任一项所述的rRNA分子,其中所述rRNA包含至少一个经修饰核苷酸。
63. 根据权利要求62所述的rRNA分子,其中所述经修饰核苷酸是N1-甲基-假尿苷。
64. 根据权利要求63所述的rRNA分子,其中所有尿苷残基是N1-甲基-假尿苷残基。
65. DNA分子,其编码根据权利要求36至64中任一项所述的rRNA分子。
66. 根据权利要求36至64中任一项所述的rRNA分子或权利要求65所述的DNA分子,其中所述rRNA或DNA分子是线性的或环状的。
67. 根据权利要求36至66中任一项所述的rRNA或DNA分子,其与脂质配制在一起。
68. 药物组合物,其包含根据权利要求36至67中任一项所述的rRNA或DNA分子以及可药用载体或赋形剂。
69. 根据权利要求36至67中任一项所述的rRNA或DNA分子或者根据权利要求68所述的药物组合物,其用于治疗。
70. 用于在对象中提高免疫应答的方法,其包括向所述对象施用编码抗原的rRNA分子,并且所述rRNA分子包含自复制RNA病毒的经修饰5' 调节区,优选甲病毒的经修饰5' 调节区,所述经修饰调节区包含位于(SEQ ID NO:2)的第67、244、245、246、248位中的一个或多个处的点突变。
71. 用于在对象中治疗癌症的方法,其包括向所述对象施用编码肿瘤抗原的rRNA分子,并且所述rRNA包含自复制RNA病毒的经修饰5' 调节区,优选甲病毒的经修饰5' 调节区,所述经修饰调节区包含位于(SEQ ID NO:2)的第67、244、245、246、248位中的一个或多个处的点突变。
72. 用于将基因功能提供给缺乏这样的基因功能的对象的方法,其包括向所述对象施用编码基因产物(治疗性蛋白质)的rRNA分子以提供所述基因功能,并且所述rRNA分子包含自复制RNA病毒的经修饰5' 调节区,优选甲病毒的经修饰5' 调节区,所述经修饰调节区包含位于(SEQ ID NO:2)的第67、244、245、246、248位中的一个或多个处的点突变。
73. 根据权利要求70至72中任一项所述的方法,其中所述rRNA分子包含位于所述5' 调节区(SEQ ID NO:2)的第67位处以及第244、245、246、248位中的一个或多个处的点突变。
74. 根据权利要求70至73中任一项所述的方法,其中所述rRNA分子是包含至少一个经修饰核苷酸的自扩增RNA分子,所述经修饰核苷酸优选N1-甲基-假尿苷。

用于确定用于提高经修饰可复制RNA功能的突变的方法以及相关组合物及其用途

技术领域

[0001] 本发明涉及用于恢复或提高含有经修饰核苷酸的可复制RNA被复制和/或翻译的能力的方法。所述方法包括鉴定核苷酸变化例如在可复制RNA中的核苷酸变化,其补偿了含有经修饰核苷酸的可复制RNA分子的复制和/或翻译能力的降低。本发明还涉及并入这样的鉴定出的核苷酸变化的含有经修饰核苷酸的可复制RNA分子,以及这样的可复制RNA分子在治疗中的用途。

背景技术

[0002] 最近,基于mRNA的疫苗在抗击Covid-19流行病的临床研究中证明了其免疫原性。这些RNA疫苗高度有效,并且诱导了非常强的T细胞免疫应答和高水平的中和抗体(Walsh et al.,2020,N Engl J Med383:2439-2450;Sahin et al.,2020,Nature 586:594-599)。获得监管批准的前两种mRNA疫苗包含经化学修饰的核苷酸,N1-甲基-假尿苷(1mψ)代替尿苷。这种修饰通过在很大程度上避免刺激导致干扰素应答的固有免疫途径来提高免疫活性细胞中mRNA的翻译(Andries et al.,2015,J Control Release 217:337-344)。

[0003] 这些批准的RNA疫苗需要每剂30至100μg RNA,并间隔数周连续两次给药(初免-加强方案(prime-boost regimen))。这最终需要60至200g RNA对100万人进行免疫接种。因此,剂量减少到小于1μg将对于向群体提供针对新的病原体的疫苗而所需的生产时间具有巨大影响。

[0004] 正在研究中的有望实现显著剂量降低的疫苗方法是使用自扩增RNA(self-amplifying RNA,saRNA)。来自甲病毒基因组的saRNA可通过用所期望的免疫应答针对的抗原替换甲病毒结构基因来改造。saRNA编码甲病毒复制酶,所述甲病毒复制酶具有复制saRNA分子的所有酶功能,因此导致输入疫苗量的扩增。遗憾的是,由saRNA引起的固有免疫应答强烈抑制了saRNA疫苗的有效性,这可能是在非人灵长类中进行的saRNA Covid-19疫苗的临床前研究中使用了相当大量的saRNA的原因(Erasmus et al.,2020,Sci Transl Med 12:eabc9396)。

[0005] 甲病毒是正链RNA病毒的典型代表。甲病毒的宿主包括广泛多种的生物体,包括昆虫、鱼和哺乳动物,例如驯养动物和人。甲病毒在受感染细胞的细胞质中复制(关于甲病毒生命周期的综述,参见José et al.,2009,Future Microbiol.4:837-856)。许多甲病毒的总基因组长度通常为11,000至12,000个核苷酸,并且基因组RNA通常具有5'帽和3' poly(A)尾。甲病毒的基因组编码非结构蛋白(参与病毒RNA的转录、修饰和复制以及蛋白质修饰)和结构蛋白(形成病毒颗粒)。基因组中通常有两个开放阅读框(open reading frame,ORF)。四种非结构蛋白(nsP1至nsP4)通常由在基因组5'末端附近开始的第一ORF一起编码,而甲病毒结构蛋白由存在于第一ORF下游并延伸在基因组3'末端附近的第二ORF一起编码。通常而言,第一ORF多于第二ORF,比例大约为2:1。

[0006] 在由甲病毒感染的细胞中,仅非结构蛋白从基因组RNA翻译,而结构蛋白可从亚基

基因组转录物翻译,所述亚基因组转录物是类似于真核信使RNA(messenger RNA,mRNA;Gould et al.,2010,Antiviral Res.87:111-124)的RNA分子。在感染之后,即在病毒生命周期的早期阶段,(+)链基因组RNA直接充当信使RNA用于翻译编码非结构多蛋白(nsP1234)的开放阅读框。在一些甲病毒中,在nsP3和nsP4的编码序列之间存在乳白终止密码子(opal stop codon):当翻译在乳白终止密码子处终止时产生含有nsP1、nsP2和nsP3的多蛋白P123以及在通读该乳白密码子之后产生另外含有nsP4的多蛋白P1234(Strauss&Strauss,1994, Microbiol.Rev.58:491-562;Rupp et al.,2015,J.Gen.Virology 96:2483-2500)。nsP1234被自体蛋白水解切割成片段nsP123和nsP4。多肽nsP123和nsP4缔合形成(-)链复制酶复合物,其使用(+)链基因组RNA作为模板转录(-)链RNA。通常在后期阶段,nsP123片段被完全切割成单独的蛋白nsP1、nsP2和nsP3(Shirako&Strauss,1994,J.Virol.68:1874-1885)。所有四种蛋白装配以形成(+)链复制酶复合体,其使用基因组RNA的(-)链互补序列作为模板合成新的(+)链基因组(Kim et al.,2004,Virology 323:153-163,Vasiljeva et al.,2003,J.Biol.Chem.278:41636-41645)。

[0007] 在受感染的细胞中,亚基因组RNA以及新的基因组RNA由nsP1提供5'帽(Pettersson et al.,1980,Eur.J.Biochem.105:435-443;Rozanov et al.,1992,J.Gen.Virology 73:2129-2134),并由nsP4提供多腺苷酸[poly(A)]尾(Rubach et al.,2009,Virology 384:201-208)。因此,亚基因组RNA和基因组RNA二者都类似于信使RNA(mRNA)。

[0008] 甲病毒结构蛋白(核心核衣壳蛋白C、包膜蛋白E2和包膜蛋白E1,病毒颗粒的所有成分)通常由亚基因组启动子控制下的一个单个开放阅读框编码(Strauss&Strauss,1994, Microbiol.Rev.58:491-562)。亚基因组启动子被顺式作用的甲病毒非结构蛋白所鉴定。特别地,甲病毒复制酶使用基因组RNA的(-)链互补序列作为模板合成(+)链亚基因组转录物。(+)链亚基因组转录物编码甲病毒结构蛋白(Kim et al.,2004,Virology323:153-163,Vasiljeva et al.,2003,J.Biol.Chem.278:41636-41645)。亚基因组RNA转录物用作用于将编码结构蛋白的开放阅读框翻译为一种多蛋白的模板,并且多蛋白被切割以产生结构蛋白。在宿主细胞中甲病毒感染的晚期,位于nsP2编码序列内的包装信号确保将基因组RNA选择性包装到由结构蛋白包装的出芽病毒粒子(budding virion)中(White et al.,1998,J.Virol.72:4320-4326)。

[0009] 在受感染的细胞中,(-)链RNA合成通常仅在感染后的前3至4小时观察到,并且在晚期不可检出,此时仅(+)链RNA(基因组和亚基因组二者的)的合成可观察到。根据Frolov et al.,2001,RNA 7:1638-1651,RNA合成调节的主要模型表明对非结构多蛋白的加工的依赖性:非结构多蛋白nsP1234的初始切割产生nsP123和nsP4;nsP4充当RNA依赖性RNA聚合酶(RNA-dependent RNApolymerase,RdRp),其对(-)链合成具有活性,但对(+)链RNA的产生的效率低。多蛋白nsP123的进一步加工,包括在nsP2/nsP3连接处的切割,改变了复制酶的模板特异性,从而增加(+)链RNA的合成并减少或终止(-)链RNA的合成。

[0010] 甲病毒RNA的合成也受顺式作用RNA元件的调节,包括四个保守序列元件(conserved sequence element,CSE;Strauss&Strauss,1994, Microbiol.Rev.58:491-562;和Frolov,2001,RNA7:1638-1651)。甲病毒基因组包含四种保守序列元件(conserved sequence element,CSE),其被认为对宿主细胞中的病毒RNA复制是重要的。在病毒基因组

的5'末端或附近存在的CSE1,被认为作为从(-)链模板合成(+)链的启动子。位于CSE 1下游但仍靠近nsP1编码序列中基因组5'末端的CSE 2,被认为充当启动从基因组RNA模板的(-)链合成的启动子(注意到不包含CSE 2的亚基因组RNA转录物不作为(-)链合成的模板)。CSE 3位于非结构蛋白和结构蛋白的编码序列之间的连接区中,并且充当用于亚基因组转录物的有效转录的核心启动子。最后,刚好位于甲病毒基因组3'非翻译区的poly(A)序列的上游的CSE 4,其被理解为作为启动(-)链合成的核心启动子(José et al., 2009, *Future Microbiol.* 4:837-856)。甲病毒的CSE 4和poly(A)尾被理解为共同起作用以进行有效的(-)链合成(Hardy&Rice, 2005, *J. Virol.* 79:4630-4639)。除甲病毒蛋白质外,宿主细胞因子(可能是蛋白质)也可与保守序列元件结合。甲病毒基因组的5'复制识别序列不仅参与翻译起始,还包含参与病毒RNA合成的两个保守序列元件CSE 1和CSE 2。对于CSE 1和2的功能,认为二级结构比线性序列更重要(Strauss&Strauss, 1994, *Microbiol. Rev.* 58:491-562)。

[0011] 已经提出了将甲病毒来源的载体用于将外来遗传信息递送到靶细胞或靶生物中。在简单的方法中,将编码甲病毒结构蛋白的开放阅读框替换为编码目的蛋白质的开放阅读框。基于甲病毒的反式复制系统依赖于两个分开的核酸分子上的甲病毒核苷酸序列元件:一个核酸分子编码病毒复制酶(通常为多蛋白nsP1234),并且另一个核酸分子能够被所述复制酶反式复制(因此称为反式复制系统)。反式复制需要在给定的宿主细胞中存在这两种核酸分子。能够被复制酶反式复制的核酸分子必须包含某些甲病毒序列元件,以允许通过甲病毒复制酶进行识别和RNA合成。

[0012] 鉴于经修饰mRNA疫苗的成功,对saRNA进行修饰以降低固有免疫应答似乎很有吸引力。然而,已经观察到经修饰核苷酸抑制了saRNA的复制和翻译(Erasmus et al., 2020, *Sci Transl Med* 12:eabc9396)。迄今为止,尚无公开RNA修饰与saRNA功能不相容原因的研究。因此,本领域仍然需要尽管含有经修饰核苷酸但具有足够的复制和/或翻译功能的saRNA分子。本发明实现了这样的需求。

发明内容

[0013] 本发明总体上涉及提高自复制RNA分子(也称为复制子或可复制RNA(replicable RNA, rRNA或saRNA))复制和/或翻译来表达编码的蛋白质的能力,所述自复制RNA分子包含除尿嘧啶、腺苷、胞嘧啶和鸟嘌呤以外的核苷酸。因此,本发明涉及用于鉴定序列变化的方法,所述序列变化恢复或提高了含有经修饰核苷酸的可复制RNA分子的降低的复制和/或翻译能力。另外,本发明涉及包含这些序列变化的可复制RNA分子,以及这样的可复制RNA分子在方法中的以下用途:用于在细胞中表达蛋白质,或者用于提高针对由可复制RNA分子编码的蛋白质的免疫应答(优选细胞毒性免疫应答),以及用于治疗或预防疾病或病症,其中这样的免疫应答导致/造成疾病或病症这样的治疗或预防。

[0014] 本发明部分地基于5'-保守序列元件(CSE)或3'-保守序列元件内的RNA二级结构在核苷酸修饰时改变其形状或稳定性的假设。RNA复制依赖于RNA模板与复制酶的相互作用。因此,RNA与复制酶的不适当相互作用将极大地影响RNA依赖性RNA转录。由于RNA结构依赖于核苷酸序列,本发明人提出,通过用经修饰核苷酸(例如N1-甲基-假尿嘧啶)替换A、C、G和U的标准核苷酸,经修饰RNA可在序列变化时调整与复制酶适当相互作用的结构。如通过本文中公开的实验结果所示出的,本发明人表明体外进化方法允许鉴定序列变化,从而允

许恢复或提高rRNA的功能。

[0015] 在一个方面中,本发明涉及用于在包含经修饰核苷酸的可复制RNA(rRNA)(经修饰rRNA)中鉴定序列变化的方法,所述经修饰核苷酸至少部分恢复或提高经修饰rRNA功能,所述方法包括(a)用编码目的基因的经修饰rRNA对表达RNA依赖性RNA聚合酶(复制酶)的细胞进行转染;以及(b)鉴定存在于表达目的基因的转染细胞中编码目的基因的rRNA分子内的序列变化。rRNA可以是单链RNA分子,其可被翻译(+链)并可编码或不可编码RNA依赖性RNA聚合酶(复制酶),但其包含这样的核苷酸序列:能够使该分子由单独提供的复制酶反式复制或者由同一rRNA编码的复制酶顺式复制。rRNA分子也可称为反式复制子(trans-replicon,TR)或复制子。另外,rRNA可包含编码基因序列(例如抗原)的密码子优化的或野生型序列,并且rRNA可包含一个或更多个针对rRNA在稳定性和翻译效率方面的最大效力而优化的结构元件(5'帽、5'UTR、3'UTR、poly(A)尾)。在一个实施方案中,序列变化存在于复制子的5'调节区中。

[0016] 在一个实施方案中,恢复或提高rRNA或者经核苷酸修饰的rRNA的功能可通过恢复或提高目的基因的表达水平来确定,例如在转染之后在固定时间段之后,例如在转染之后在1、2、3、4、5、6、12、18、24小时或者1、2、3、4或5天之后。在一个实施方案中,恢复或提高rRNA的功能可通过恢复或提高rRNA复制的能力来确定,如例如通过在转染之后在固定时间段例如1、2、3、4、5、6、12、18、24小时或者1、2、3、4或5天之后在经转染细胞中rRNA拷贝的总数目来测量。

[0017] 在一个实施方案中,可相对于步骤(a)中转染的经修饰rRNA的序列来确定序列变化。还可相对于参考序列来确定序列变化。在一个实施方案中,该方法的步骤(a)还可包括在转染之后,从表达目的基因的经转染细胞中将编码目的基因的rRNA分子进行分离,以提供分离的rRNA分子。在一个实施方案中,步骤(a)还可包括用编码目的基因并包含经修饰核苷酸的rRNA分子的后代对表达复制酶的细胞进行转染的步骤,所述后代可通过复制分离的rRNA分子来提供。

[0018] 在一个实施方案中,该方法还可包括重复以下步骤:从表达目的基因的经转染细胞中将编码目的基因的rRNA分子进行分离提供分离的rRNA分子,并且用编码目的基因并包含经修饰核苷酸的rRNA分子的后代对表达复制酶的细胞进行转染,所述后代可通过复制分离的rRNA分子来提供。这些步骤可重复数(n)次,其中n可以是至少为1的整数。这些步骤可重复1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20次或更多次。在一个实施方案中,复制分离的rRNA分子包括(i)将编码目的基因的分离的rRNA分子进行逆转录,以形成能够被体外转录的DNA分子;以及(ii)在相同的经修饰核苷酸的存在下体外转录DNA分子,以产生包含经修饰核苷酸并编码目的基因的经修饰rRNA分子的后代。逆转录可通过包括一步法逆转录PCR(reverse transcription PCR,RT-PCR)的方法来完成。

[0019] 在一个实施方案中,该方法还可包括以下步骤:(c)将至少一个鉴定出的序列变化并入到步骤(a)的经修饰rRNA中;(d)用步骤(c)中产生的经修饰rRNA对表达复制酶的细胞进行转染;以及(e)鉴定从步骤(d)中产生的表达目的基因的经转染细胞中分离的rRNA分子内的序列变化。在一个实施方案中,该方法中使用的细胞是没有干扰素应答的细胞。

[0020] 在一些实施方案中,可复制RNA(rRNA)中的经修饰核苷酸不是5'帽结构或者是除5'帽结构以外的。在一个实施方案中,可用于该方法的经修饰rRNA中的经修饰核苷酸的数

目可以是rRNA中核苷酸总数目的至少10%、20%、30%、40%、50%、60%、70%、80%、90%或100%。优选地,可用于该方法的经修饰rRNA中的经修饰核苷酸的数目可以是rRNA中核苷酸总数目的至少30%。更优选地,可用于该方法的经修饰rRNA中的经修饰核苷酸的数目可以是rRNA中核苷酸总数目的至少50%。经修饰核苷酸可以是经修饰的尿苷残基、经修饰的腺嘌呤残基、经修饰的鸟嘌呤残基、经修饰的胞嘧啶残基或者上述的两种或更多种的任意组合。在一个实施方案中,经修饰核苷酸是经修饰的尿苷,例如假尿苷或N1-甲基-假尿苷。在一个实施方案中,体外转录的经修饰rRNA中经修饰的尿苷残基的数目可以是尿苷残基总数目的至少10%、20%、30%、40%、50%、60%、70%、80%、90%或100%。体外转录的经修饰rRNA中经修饰的尿苷残基(例如,假尿苷残基或N1-甲基-假尿苷残基)的数目可以是尿苷残基总数目的至少30%。优选地,体外转录的经修饰rRNA中经修饰的尿苷残基(例如,假尿苷残基或N1-甲基-假尿苷残基)的数目可以是尿苷残基总数目的至少50%。体外转录的经修饰rRNA中经修饰的尿苷残基(例如,假尿苷残基或N1-甲基-假尿苷残基)的数目可以是尿苷残基总数目的约100%(即基本上全部)。在一个实施方案中,体外转录的经修饰rRNA中N1-甲基-假尿苷残基的数目可以是尿苷残基总数目的至少10%、20%、30%、40%、50%、60%、70%、80%、90%或100%。优选地,体外转录的经修饰rRNA中N1-甲基-假尿苷残基的数目可以是尿苷残基总数目的至少30%。更优选地,体外转录的经修饰rRNA中N1-甲基-假尿苷残基的数目可以是尿苷残基总数目的至少50%。

[0021] 在一个实施方案中,可用于该方法的rRNA基本上是单链的、正链的RNA病毒的基因组,所述RNA病毒例如甲病毒、小RNA病毒、黄病毒,任选地所述RNA病毒不编码该病毒的功能性结构蛋白。

[0022] 在一个实施方案中,可用于该方法的复制酶可来源于自复制的RNA病毒,例如甲病毒,例如塞姆利基森林病毒(Semliki Forest virus,SFV)、委内瑞拉马脑炎病毒(Venezuelan equine encephalitis virus,VEEV)、辛德比斯病毒(Sindbis virus)、东部马脑炎病毒(Eastern equine encephalitis virus,EEEV)、西部马脑炎病毒或基孔肯亚病毒(Chikungunya virus)。在一个实施方案中,甲病毒可以是SFV或VEEV或EEEV或辛德比斯病毒。

[0023] 在一个实施方案中,可用于该方法的rRNA包含不具有起始密码子的5'调节区。在一个实施方案中,复制酶和rRNA中的可用于复制酶复制所需的方法中的调节序列来源于相同的甲病毒或来源于不同的甲病毒。

[0024] 在一个实施方案中,其中rRNA还包含编码复制酶的开放阅读框,通过将复制酶开放阅读框的翻译置于内部核糖体进入位点(internal ribosome entry site,IRES)的翻译控制下,使复制酶开放阅读框的翻译与5'端帽解偶联。在一个实施方案中,rRNA可以是未加帽的。在一个实施方案中,rRNA可具有用于表达在IRES上游的另外基因的开放阅读框。

[0025] 在一个实施方案中,编码的目的基因可以是抗原或报道基因,例如细胞表面表达的蛋白质或萤光素酶或GFP。

[0026] 在一个实施方案中,可用于该方法的细胞中复制酶的表达可通过瞬时表达或通过组成型表达进行。在一个实施方案中,在rRNA(不一定编码复制酶)的转染之前、同时或之后,用编码复制酶的核酸分子转染细胞。在一个实施方案中,编码复制酶的核酸分子是RNA分子,例如mRNA分子。在一个实施方案中,该方法的步骤(a)包括用编码目的基因的自扩增

RNA转染细胞,该自复制RNA包含至少一个经修饰核苷酸。

[0027] 在一个实施方案中,部分恢复或提高的rRNA的功能可以是由复制酶复制的能力和/或翻译的能力;例如,如由5' 调节区所指导的。在一个实施方案中,功能可恢复了至少10%或者可提高了至少10%。

[0028] 在一个实施方案中,鉴定序列变化的步骤可包括对rRNA分子的至少一部分进行测序。在一个实施方案中,测序的rRNA部分可以是5' 调节区和/或测序的rRNA部分是3' 调节区。在一个实施方案中,可通过分析由测序(例如下一代测序)产生的序列特异性读数频率来确定序列变化。

[0029] 在一个实施方案中,可进行该方法以提高包含至少一个经修饰核苷酸的自扩增RNA的复制效率。

[0030] 在一个实施方案中,该方法还包括通过并入至少一个鉴定出的序列变化来修饰包含相同经修饰核苷酸的rRNA或自复制RNA的核苷酸序列,所述序列变化部分恢复或提高包含经修饰核苷酸的rRNA或自复制RNA的功能。

[0031] 在一个方面中,本发明涉及通过上述本发明的方法获得的可复制RNA(rRNA)分子,该分子包含一个或更多个经修饰核苷酸。在一个实施方案中,获得的rRNA分子包含自复制RNA病毒的突变的5' 调节区。在一个实施方案中,突变的5' 调节区包含一个或更多个点突变。在一个特定实施方案中,通过本发明的方法所获得的包含一个或更多个经修饰核苷酸的rRNA分子与未经受该方法的亲本rRNA分子相比,能够以水平的提高由复制酶复制/转录。在一个实施方案中,rRNA分子包含rRNA分子的总核苷酸中至少30%、50%、75%的经修饰核苷酸。因此,在一些实施方案中,本发明提供了可通过如本文中所述的本发明的方法所获得的可复制RNA(rRNA)分子,其中rRNA分子中尿苷残基总数目的至少50%(优选至少75%,最优选约100%)是经修饰的尿苷残基(例如假尿苷残基或N1-甲基-假尿苷残基),并且其中rRNA分子能够被复制酶以与未经受该方法的亲本rRNA分子相比相似或提高的水平进行复制/转录。在一些实施方案中,与包含rRNA分子的野生型5' 调节区的相应rRNA分子相比,任选地与具有包含SEQ ID NO:1或SEQ ID NO:2的5' 调节区(VEEV特立尼达驴株(VEEV Trinidad donkey strain)(登记号L01442)的5' 调节区的野生型或经修饰形式)的相应rRNA分子相比,该rRNA分子可能以相似或提高的水平由复制酶复制/转录。复制酶可以是甲病毒复制酶。在一些实施方案中,复制酶包含来自以下的非结构蛋白:塞姆利基森林病毒(SFV)、辛德比斯病毒、委内瑞拉马脑炎病毒(VEEV)、东部马脑炎病毒(EEEV)或基孔肯亚病毒(chikungunya virus,CHIKV),优选SFV、VEEV或EEV,更优选SFV。

[0032] 在本发明的一个方面中,经修饰的突变rRNA分子包含自复制RNA病毒的经修饰5' 调节区,该经修饰调节区可包含在5' 调节区(SEQ ID NO:2)的第67、244、245、246、248位中的一个或更多个处的点突变,或者对应于SEQ ID NO:2中那些位置的甲病毒5' 调节区中的位置的点突变。

[0033] 在一个实施方案中,rRNA分子可包含至少一个经修饰核苷酸或核苷酸。

[0034] SEQ ID NO:2是VEEV特立尼达驴株(登录号L01442)的5' 调节序列,其与野生型序列(SEQ ID NO:1)的5' 末端相比尤其在病毒5' 末端通过添加G残基来修饰。RNA的DNA模板中的这种修饰允许使用T7聚合酶进行有效的体外转录。

[0035] 在一个实施方案中,rRNA分子在5' 调节区中还可包含在5' 调节区(SEQ ID NO:2)

的第4位处的点突变或者在对应于SEQ ID NO:2中该位置的甲病毒5' 调节区中的位置处的点突变,所述点突变优选地是用腺苷残基替换鸟嘌呤残基(G4A)。

[0036] 在一个实施方案中,rRNA分子在5' 调节区中不包含在5' 调节区(SEQ ID NO:2)的第4位处的点突变或者在对应于SEQ ID NO:2中该位置的甲病毒5' 调节区中的位置处的点突变,所述点突变优选地不是用腺苷残基替换鸟嘌呤残基(G4A)。

[0037] 自复制RNA病毒可以是甲病毒。

[0038] SEQ ID NO:1是VEEV特立尼达驴株(登录号L01442)的野生型5' 调节区,其不包含5' 末端(most)G残基。因此,SEQ ID NO:2中的第4、67、244、245、246和248位分别对应于SEQ ID NO:1中的第3、66、243、244、245和247位,并且SEQ ID NO:2中的第4、67、244、245、246和248位处核苷酸的同一性分别与SEQ ID NO:1中的第3、66、243、244、245和247位处核苷酸的同一性相同。

[0039] 在多个实施方案中,点突变可位于SEQ ID NO:2的第67位处,第67和244位处,第67和246位处,第67和248位处,第67、245和248位处,第4和67位处,第4、67和244位处,第4、67和246位处,第4、67和248位处,第4、67、245和248位处,或者其任意组合处,或者对应于SEQ ID NO:2中这样的位置的甲病毒中的位置处。

[0040] 在多个实施方案中,点突变可位于SEQ ID NO:2的以下位置中一个或更多个或者组合的位置处,或者在甲病毒的5' 调节区中相应的位置处:67,244,245,246,248,67和244,67和245,67和246,67和248,244和245,244和246,244和248,245和246,245和248,246和248,67、244和245,67、244和246,67、244和248,67、245和246,67、245和248,67、246和248,244、245和246,244、245和248,244、246和248,245、246和248,67、244、245和246,67、244、245和248,67、244、246和248,67、245、246和248,244、245、246和248,67、244、245、246和248。

[0041] 在一个实施方案中,rRNA分子的经修饰调节区包含这样的点突变,所述点突变在第67位处或在对应于SEQ ID NO:2中该位置的甲病毒5' 调节区中的位置处以及在5' 调节区(SEQ ID NO:2)的第244、245、246、248位中的一个或更多个处或者在对应于SEQ ID NO:2中那些位置的甲病毒5' 调节区中的位置处。

[0042] 在每个前述位置或位置的组合中,点突变还可位于SEQ ID NO:2的第4位处或在甲病毒的5' 调节区中相应的位置处。

[0043] 在一个实施方案中,在不同位置处的点突变可以是G4A、A67C、G244A、C245A、G246A或C248A。

[0044] 在一个实施方案中,自复制RNA病毒可以是SFV或VEEV或EEEV或辛德比斯病毒。

[0045] 在一个实施方案中,另一种甲病毒中的相应位置可以是相对于5' 调节区第一个核苷酸的相同位置和/或在另一种甲病毒5' 调节区中存在的茎/环结构的一者中存在的相同的相对位置。

[0046] 在一个实施方案中,rRNA分子还可包含一个或更多个编码区,例如包含编码目的基因的序列的一个或更多个编码区。例如,编码的目的基因可以是抗原或治疗性蛋白或核酸或报道基因。在一个实施方案中,抗原是肿瘤、病毒、细菌或真菌抗原,或者变应原。

[0047] 在一个实施方案中,rRNA分子可编码RNA依赖性RNA聚合酶(复制酶),例如甲病毒复制酶。在一个实施方案中,rRNA分子可包含5' 调节区和复制酶的编码序列,其中5' 调节区

和编码的复制酶来源于相同的自复制RNA病毒,例如来源于相同的甲病毒。在一个实施方案中,rRNA分子可包含5'调节区和复制酶的编码序列,其中5'调节区和编码的复制酶来源于不同的自复制RNA病毒,例如不同的甲病毒。

[0048] 在一个实施方案中,rRNA分子可以是自扩增RNA,即编码复制酶是自扩增RNA。在一个实施方案中,rRNA分子可以是体外转录的rRNA,其在体外转录反应中产生,即不在细胞中产生。

[0049] 在一个实施方案中,其中rRNA还包含编码复制酶的开放阅读框,通过将复制酶开放阅读框的翻译置于内部核糖体进入位点(IRES)的翻译控制下,复制酶开放阅读框的翻译可与5'端帽解偶联。在一个实施方案中,rRNA可以是未加帽的。在一个实施方案中,rRNA可具有用于表达在IRES上游的另外基因的开放阅读框。

[0050] 在一个实施方案中,rRNA分子包含至少一个经修饰核苷酸,即不是腺苷、鸟嘌呤、胞嘧啶或尿嘧啶的核苷酸,例如如上所述。如本文中使用的,经修饰核苷酸可在骨架键联、糖部分或碱基中被修饰,并包括经修饰核苷。在一个实施方案中,经修饰核苷酸可以是N1-甲基-假尿苷。在一个实施方案中,rRNA的所有尿苷残基均是N1-甲基-假尿苷残基。

[0051] 在一个实施方案中,rRNA可以是稳定的rRNA。在一个实施方案中,rRNA可包含替换了至少一个尿苷的经修饰核苷。在一个实施方案中,rRNA可包含替换了每个尿苷的经修饰核苷。在一个实施方案中,经修饰核苷可独立地选自假尿苷(ψ)、N1-甲基-假尿苷(m1 ψ 或1mY)和5-甲基-尿苷(m5U)。

[0052] 在一个方面中,本发明涉及包含SEQ ID NO:2中所述序列的RNA分子,其中第4、67、244、245、246和248位中的一个或多个具有点突变。在多个实施方案中,点突变可在第67,244,245,246,248,67和244,67和245,67和246,67和248,244和245,244和246,244和248,245和246,245和248,246和248,67、244和245,67、244和246,67、244和248,67、245和246,67、245和248,67、246和248,244、245和246,244、245和248,244、246和248,245、246和248,67、244、245和246,67、244、245和248,67、244、246和248,67、245、246和248,244、245、246和248,67、244、245、246和248位处。在一个实施方案中,RNA分子包含SEQ ID NO:2中所述的序列,其中第67位以及第4、244、245、246和248位中的一个或多个具有点突变。在每个前述位置或位置的组合中,另外的点突变可在第4位处。在一个实施方案中,点突变可以是G4A、A67C、G244A、C245A、G246A和/或C248A。

[0053] 在一个方面中,本发明涉及编码rRNA分子的DNA分子,其包含SEQ ID NO:2本文中公开位置处的一个或多个点突变。在一个实施方案中,rRNA分子或DNA分子可以是线性的或环状的。

[0054] 在一个实施方案中,可以用能够与rRNA或DNA分子形成颗粒的试剂配制rRNA或DNA分子,例如,该试剂可以是脂质或聚亚烷基胺。在多个实施方案中,脂质可包含阳离子首基,和/或脂质可以是pH响应性脂质,和/或脂质可以是聚乙二醇化脂质(PEGylated-lipid)。在一个实施方案中,试剂可与聚肌氨酸缀合。在一个实施方案中,由rRNA或DNA分子以及试剂形成的颗粒可以是基于聚合物的多聚复合物(polyplex,PLX)或脂质纳米粒(lipid nanoparticle,LNP),其中LNP优选地是脂质复合物(lipoplex,LPX)或脂质体。在一个实施方案中,颗粒还可包含至少一个磷脂酰丝氨酸。在一个实施方案中,颗粒可以是纳米粒,其中(i)纳米粒中正电荷的数目不超过纳米粒中负电荷的数目以及/或者(ii)纳米粒具

有中性或净负电荷以及/或者(iii)纳米粒中正电荷与负电荷的电荷比为1.4:1或更小以及/或者(iv)纳米粒的 ζ 电势为0或更小。在一个实施方案中,纳米粒中正电荷与负电荷的电荷比可以是1.4:1至1:8,优选1.2:1至1:4。

[0055] 在一个实施方案中,rRNA或DNA可配制或待配制成液体、固体或其组合。在一个实施方案中,rRNA或DNA可配制成或待配制成用于注射。在一个实施方案中,rRNA或DNA可配制或待配制成用于肌肉施用。在一个实施方案中,rRNA或DNA可配制或待配制成颗粒。在一个实施方案中,颗粒是脂质纳米粒(LNP)或脂质复合体(LPX)颗粒。在一个实施方案中,LNP颗粒包含((4-羟基丁基)氮烷二基)双(己烷-6,1-二基)双(2-己基癸酸酯)、2-[(聚乙二醇)-2000]-N,N-双十四烷基乙酰胺、1,2-二硬脂酰基-sn-甘油基-3-磷酸胆碱和胆固醇。

[0056] 在一个实施方案中,可通过将rRNA与脂质体混合来获得rRNA脂质复合物颗粒。在一个实施方案中,可通过将rRNA与脂质混合来获得rRNA脂质复合物颗粒。

[0057] 在一个实施方案中,rRNA可配制或待配制成胶体。在一个实施方案中,rRNA可配制或待配制成颗粒,以形成胶体的分散相。在一个实施方案中,rRNA中的50%或更多、75%或更多或者85%或更多存在于分散相中。在一个实施方案中,rRNA可配制或待配制成包含rRNA和脂质的颗粒。在一个实施方案中,可通过将溶解于水相的rRNA与溶解于有机相的脂质接触来形成颗粒。在一个实施方案中,有机相可包含乙醇。在一个实施方案中,可通过将溶解于水相中的rRNA与分散于水相中的脂质接触来形成颗粒。在一个实施方案中,分散于水相中的脂质形成脂质体。

[0058] 在一个方面中,本发明涉及药物组合物,其包含含有本文中所述点突变的rRNA分子或编码这样的rRNA分子的DNA分子以及可药用载体或赋形剂。

[0059] 在一个方面中,本发明涉及在细胞中表达目的基因产物的方法,所述方法包括向细胞提供rRNA,所述rRNA包含自复制RNA病毒(优选甲病毒)的经修饰5'调节区和编码目的基因产物的开放阅读框,该经修饰调节区包含位于SEQ ID NO:2的第67、244、245、246、248位中的一个或多个处的点突变,优选地位于SEQ ID NO:2的第67位处以及第244、245、246、248位中的一个或多个处的点突变。

[0060] 在一个方面中,本发明涉及可用于治疗的包含本文中所述点突变的rRNA或编码DNA分子或包含该rRNA或DNA分子的药物组合物。在一个方面中,本发明涉及用于在对象中提高免疫应答的方法,其包括向对象施用编码抗原的rRNA分子,并且所述rRNA分子包含自复制RNA病毒(优选甲病毒)的经修饰5'调节区,所述经修饰调节区包含位于SEQ ID NO:2的第67、244、245、246、248位中的一个或多个处的点突变,优选位于SEQ ID NO:2的第67位处以及第244、245、246、248位中的一个或多个处的点突变。

[0061] 在一个方面中,本发明涉及用于在对象中治疗癌症的方法,其包括向对象施用编码肿瘤抗原的rRNA分子,并且所述rRNA包含自复制RNA病毒(优选甲病毒)的经修饰5'调节区,所述经修饰调节区包含位于SEQ ID NO:2的第67、244、245、246、248位中的一个或多个处的点突变,优选地位于SEQ ID NO:2的第67位处以及第244、245、246、248位中的一个或多个处的点突变。

[0062] 在一个方面中,本发明涉及用于将基因功能提供给缺乏这样基因功能的对象的方法,其包括向对象施用编码基因产物(治疗性蛋白)的rRNA分子以提供基因功能,并且所述rRNA分子包含自复制RNA病毒(优选甲病毒)的经修饰5'调节区,所述经修饰调节区包含位

于SEQ ID NO:2的第67、244、245、246、248位中的一个或更多个处的点突变,优选地位于SEQ ID NO:2的第67位处以及第244、245、246、248位中的一个或更多个处的点突变。

[0063] 在一个实施方案中,点突变还可位于SEQ ID NO:2的第4位处。在一个实施方案中,可用于表达基因产物或可用于治疗的rRNA分子包含至少一个经修饰核苷酸,例如N1-甲基-假尿苷。这样的rRNA可以是自扩增RNA分子,因为其还包含编码复制酶的序列。

[0064] rRNA分子的任何前述特征均可同样适用于这样的rRNA:在用于鉴定用于恢复或提高rRNA功能的序列变化的本发明的方法中使用的rRNA,以及将具有这样的鉴定出的序列变化并入其序列的rRNA分子中使用的rRNA。

具体实施方式

[0065] 尽管下面详细描述了本发明,但应理解本发明不限于本文中所述的特定方法、方案和试剂,因为这些可变化。还应理解,本文中使用的术语仅用于描述特定实施方案的目的,并不旨在限制本发明的范围,本发明的范围仅由所附权利要求书限制。除非另有限定,否则本文中使用的所有技术和科学术语具有与本领域普通技术人员通常理解的含义相同的含义。

[0066] 优选地,本文中使用的术语如在“*A multilingual glossary of biotechnological terms: (IUPAC Recommendations)*”, H.G.W. Leuenberger, B. Nagel, and H. Kölbl, Eds., Helvetica Chimica Acta, CH-4010 Basel, Switzerland, (1995) 中所述定义。

[0067] 除非另有说明,否则本发明的实践将采用化学、生物化学、细胞生物学、免疫学和重组DNA技术的常规方法,这些方法在本领域的文献中有解释(参见,例如, *Molecular Cloning: A Laboratory Manual*, 2nd Edition, J. Sambrook et al. eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989)。

[0068] 在下文中,将描述本发明的一些要素。这些要素与一些特定实施方案一起列出,然而,应理解,其可以以任何方式和任何数量组合以产生另外的实施方案。不应将多个描述的实施例和优选实施方案解释为将本发明仅限于明确描述的实施方案。该描述应被理解为公开并涵盖将明确描述的实施方案与任何数量的所公开和/或优选的要素组合的实施方案。此外,除非上下文另有说明,否则本申请中所有描述的元素任何排列和组合都应被认为通过本说明书公开。

[0069] 术语“约”意指大约或接近,并且在本文中所述的数值或范围的上下文中,优选意指所列举或要求保护的数值或范围的 $\pm 10\%$ 。

[0070] 除非在本文中另有说明或与上下文明显矛盾,否则在描述本发明的上下文中(特别是在权利要求的上下文中)使用的未限定数量的术语以及类似的引用应被解释为涵盖单数和复数两者。本文中对数值范围的叙述仅旨在用作单独的提及落入该范围内的每个单独值的简写方法。除非本文中另有说明,否则将每个单独的值并入本说明书中,如同其在本文中单独引用一样。除非本文中另有说明或与上下文明显矛盾,否则本文中所述的所有方法均可以以任何合适的顺序进行。本文中提供的任何和所有实施例或示例性语言(例如,“例如”)的使用仅旨在更好地说明本发明,而不是对本发明的范围进行限制,除非另外声明。说明书中的任何语言都不应被解释为表示任何对于本发明的实践必不可少的未要求保护的

要素。

[0071] 除非另有明确说明,否则在本文件的上下文中使用术语“包含”来表示除了由“包含”引入的列表的成员之外可任选地存在其他成员。然而,预期作为本发明的一个特定实施方案,术语“包含”包括不存在其他成员的可能性,即,为了本实施方案的目的,“包含”应理解为具有“由...组成”的含义。

[0072] 由通用术语表征的组分的相对量的指示意指所述通用术语涵盖的所有特定变体或成员的总量。如果指定由通用术语定义的某个组分以某个相对量存在,并且如果该组分被进一步表征为通用术语所涵盖的特定变体或成员,则意指没有另外存在由通用术语涵盖的其他变体或成员,使得通用术语所涵盖的组分的总相对量超过指定的相对量;更优选地,根本不存在由通用术语所涵盖的其他变体或成员。

[0073] 在本说明书的全文中引用了若干文献。本文中引用的每篇文献(包括所有专利、专利申请、科学出版物、制造商的说明书、说明书等),无论是上文中还是下文中,均通过引用整体在此并入。本文中的任何内容均不应被解释为承认本发明无权先于此类公开内容。

[0074] 本文中使用的术语例如“降低”或“抑制”意指引起在水平上总体降低优选5%或更高、10%或更高、20%或更高、更优选50%或更高以及最优选75%或更高的能力。术语“抑制”或类似短语包括完全或基本上完全的抑制,即降低至零或基本上降低至零。

[0075] 术语例如“提高”或“增强”优选涉及提高或增强约至少10%、优选至少20%、优选至少30%、更优选至少40%、更优选至少50%、甚至更优选至少80%、以及最优选至少100%。

[0076] 术语“净电荷”是指整个物体(例如化合物或颗粒)上的电荷。

[0077] 具有总净正电荷的离子是阳离子,而具有总净负电荷的离子是阴离子。因此,根据本发明,阴离子是比质子具有更多电子的离子,使其具有净负电荷;以及阳离子是比质子具有更少电子的离子,使其具有净正电荷。

[0078] 关于给定的化合物或颗粒,术语如“带电荷的”,“净电荷”,“带负电荷的”或“带正电荷的”是指给定的化合物或颗粒在pH 7.0下溶解或混悬于水中时的净电荷。

[0079] 根据本发明的术语“核酸”还包括核酸在核苷酸碱基上、糖上或磷酸酯上的化学衍生化,以及含有非天然的核苷酸和核苷酸类似物的核酸。在一些实施方案中,核酸是脱氧核糖核酸(deoxyribonucleic acid, DNA)或核糖核酸(ribonucleic acid, RNA)。一般而言,核酸分子或核酸序列是指优选为脱氧核糖核酸(DNA)或核糖核酸(RNA)的核酸。根据本发明,核酸包括基因组DNA、cDNA、mRNA、病毒RNA、重组制备和化学合成的分子。根据本发明,核酸可以是单链或双链以及线性或共价闭合的环状分子的形式。

[0080] 根据本发明,“核酸序列”是指核酸(例如核糖核酸(RNA)或脱氧核糖核酸(DNA))中核苷酸的序列。该术语可指整个核酸分子(例如整个核酸分子的单链)或其一部分(例如片段)。

[0081] 根据本发明,术语“RNA”或“RNA分子”涉及包含核糖核苷酸残基并且优选完全或基本上由核糖核苷酸残基构成的分子。术语“核糖核苷酸”是指在 β -D-呋喃核糖基的2'-位处具有羟基的核苷酸。术语“RNA”包括双链RNA、单链RNA、分离的RNA,例如部分或完全纯化的RNA、基本上纯的RNA、合成RNA,以及重组产生的RNA,例如通过添加、缺失、替换和/或改变一个或更多个核苷酸而不同于天然存在的RNA的经修饰RNA。这样的改变可包括将非核苷酸物

质添加到例如RNA的末端或内部,例如在RNA的一个或多个核苷酸处。RNA分子中的核苷酸还可包含非标准核苷酸,例如非天然存在的核苷酸或化学合成的核苷酸或脱氧核苷酸。这些改变的RNA可称为类似物,特别是天然存在的RNA的类似物。

[0082] 根据本发明,RNA可以是单链或双链的。在本发明的一些实施方案中,单链RNA是优选的。术语“单链RNA”通常是指没有互补核酸分子(通常没有互补RNA分子)与其缔合的RNA分子。单链RNA可包含自补序列,其允许RNA的一部分折叠回来并形成二级结构基序,包括但不限于碱基对、茎、茎环和凸起。单链RNA可作为负链[(-)链]或作为正链[(+)链]存在。(+)链是包含或编码遗传信息的链。遗传信息可以是例如编码蛋白质的多核苷酸序列。当(+)链RNA编码蛋白质时,(+)链可直接用作翻译(蛋白质合成)的模板。(-)链是(+)链的互补链。在双链RNA的情况下,(+)链和(-)链是两个独立的RNA分子,并且这两个RNA分子彼此缔合以形成双链RNA(“双链体RNA”)。

[0083] 术语RNA的“稳定性”涉及RNA的“半衰期”。“半衰期”涉及消除分子的活性、量或数量的一半所需的时间段。在本发明的上下文中,RNA的半衰期指示所述RNA的稳定性。RNA的半衰期可能会影响RNA的“表达持续时间”。可预期,半衰期长的RNA将在延长的时间内表达。

[0084] 术语“翻译效率”涉及在特定时间段内由RNA分子提供的翻译产物的量。

[0085] 关于核酸序列,“片段”涉及核酸序列的一部分,即代表在5'和/或3'末端缩短的核酸序列的序列。优选地,核酸序列的片段包含来自所述核酸序列的至少80%、优选至少90%、95%、96%、97%、98%或99%的核苷酸残基。在本发明中,优选的是RNA分子的保留RNA稳定性和/或翻译效率的那些片段。

[0086] 关于氨基酸序列(肽或蛋白质),“片段”涉及氨基酸序列的一部分,即代表在N端和/或C端缩短的氨基酸序列的序列。可例如通过翻译缺少开放阅读框的3'末端的截短的开放阅读框来获得在C端缩短的片段(N端片段)。可例如通过翻译缺少开放阅读框的5'末端的截短的开放阅读框来获得在N端缩短的片段(C端片段),只要截短的开放阅读框包含用于开始翻译的起始密码子即可。氨基酸序列的片段包含来自氨基酸序列的氨基酸残基的例如至少1%、至少2%、至少3%、至少4%、至少5%、至少10%、至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%。

[0087] 关于例如根据本发明的核酸和氨基酸序列,术语“变体”包括任何变体,特别是突变体、病毒株变体、剪接变体、构象、同种型、等位基因变体、物种变体和物种同源物,特别是天然存在的那些。等位基因变体涉及基因正常序列的改变,其重要性通常不清楚。完整的基因测序通常鉴定给定基因的许多等位基因变体。关于核酸分子,术语“变体”包括简并核酸序列,其中根据本发明的简并核酸是由于遗传密码的简并性而在密码子序列上不同于参考核酸的核酸。物种同源物是与给定核酸或氨基酸序列具有不同起源物种的核酸或氨基酸序列。病毒同源物是与给定核酸或氨基酸序列具有不同起源病毒的核酸或氨基酸序列。

[0088] 根据本发明,与参考核酸相比,核酸变体包括单个或多个核苷酸的缺失、添加、突变、替换和/或插入。缺失包括从参考核酸中去除一个或多个核苷酸。添加变体包括一个或多个核苷酸例如1、2、3、5、10、20、30、50或多个核苷酸的5'和/或3'端融合。在替换的情况下,序列中的至少一个核苷酸被去除,并且至少一个其他核苷酸被插入其位置(例如颠换和转换)。突变包括无碱基位点、交联位点和化学改变或经修饰的碱基。插入包括向参考核酸中添加至少一个核苷酸。

[0089] 根据本发明,“核苷酸变化”可指与参考核酸相比单个或更多个核苷酸的缺失、添加、突变、替换和/或插入。在一些实施方案中,“核苷酸变化”选自与参考核酸相比单个核苷酸的缺失、单个核苷酸的添加、单个核苷酸的突变、单个核苷酸的替换和/或单个核苷酸的插入。根据本发明,与参考核酸相比,核酸变体可包含一个或更多个核苷酸变化。

[0090] 特定核酸序列的变体优选具有所述特定序列的至少一种功能性质,并且优选在功能上等同于所述特定序列,例如,具有与特定核酸序列相同或相似的特性的核酸序列。

[0091] 如下所述,本发明的一些实施方案的特征尤其在于与另一些核酸序列同源的核酸序列。这些同源序列是另一些核酸序列的变体。

[0092] 给定核酸序列与作为所述给定核酸序列的变体的核酸序列之间的同一性程度优选为至少70%,优选至少75%,优选至少80%,更优选至少85%,甚至更优选至少90%或最优选至少95%、96%、97%、98%或99%。优选给出至少约30、至少约50、至少约70、至少约90、至少约100、至少约150、至少约200、至少约250、至少约300或至少约400个核苷酸区域的同一性程度。在一些优选实施方案中,给出了参考核酸序列的整个长度的同一性程度。

[0093] “序列相似性”表示相同或代表保守氨基酸替换的氨基酸的百分比。两个多肽或核酸序列之间的“序列同一性”表示序列之间相同的氨基酸或核苷酸的百分比。

[0094] 术语“%同一性”特别地旨在是指待比较的两个序列之间在最佳比对中相同的核苷酸的百分比,所述百分比是纯统计学的,并且两个序列之间的差异可随机分布在序列的整个长度上,并且待比较序列与参考序列相比可包含添加或缺失,以获得两个序列之间的最佳比对。两个序列的比较通常在关于片段或“比较窗口”最佳比对以鉴定相应序列的局部区域之后通过比较所述序列来进行。用于比较的最佳比对可手动进行,或者借助于Smith and Waterman,1981,Ads App.Math.2:482的局部同源性算法,借助于Needleman and Wunsch,1970,J.Mol.Biol.48:443的局部同源性算法,以及借助于Pearson and Lipman,1988,Proc.Natl Acad.Sci.USA 85:2444的相似性搜索算法,或借助于使用所述算法的计算机程序(Wisconsin Genetics Software Package中的GAP、BESTFIT、FASTA、BLAST P、BLAST N和TFASTA,Genetics Computer Group,575Science Drive,Madison,Wis)。

[0095] 通过确定待比较的序列所对应的相同位置的数目,用该数目除以比较的位置数目,并且将此结果乘以100来得出百分比同一性。

[0096] 例如,可使用BLAST程序“BLAST 2sequences”,其可在网站<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>上获得。

[0097] 如果核酸与另一个核酸彼此互补,则两个序列“能够杂交”或“杂交”。如果核酸与另一核酸能够彼此形成稳定的双链体,则这两个序列“互补”。根据本发明,杂交优选在允许多核苷酸之间特异性杂交的条件(严格条件)下进行。严格条件在例如以下中进行了描述: Molecular Cloning: A Laboratory Manual, J. Sambrook et al., Editors, 2nd Edition, Cold Spring Harbor Laboratory press, Cold Spring Harbor, New York, 1989 or Current Protocols in Molecular Biology, F. M. Ausubel et al., Editors, John Wiley & Sons, Inc., New York, 并且是指例如在杂交缓冲液(3.5×SSC、0.02% Ficoll、0.02% 聚乙烯吡咯烷酮、0.02% 牛血清白蛋白、2.5mM NaH₂PO₄(pH 7)、0.5% SDS、2mM EDTA)中在65℃下杂交。SSC为0.15M氯化钠/0.15M柠檬酸钠,pH 7。杂交之后,洗涤转移了DNA的膜,例如在室温下在2×SSC中,并随后在高至68℃的温度下在0.1至0.5×SSC/0.1×SDS中。

[0098] 百分比互补性表示核酸分子中可与第二核酸序列形成氢键(例如,Watson-Crick 碱基配对)的连续残基的百分比(例如,10个中的5、6、7、8、9、10为50%、60%、70%、80%、90%和100%互补)。“完美互补”或“完全互补”意指核酸序列的所有连续残基与第二核酸序列中相同数目的连续残基氢键键合。优选地,根据本发明的互补性程度为至少70%,优选至少75%,优选至少80%,更优选至少85%,甚至更优选至少90%或最优选至少95%、96%、97%、98%或99%。最优选地,根据本发明的互补性程度是100%。

[0099] 术语“衍生物”包括核酸在核苷酸碱基上、糖上或磷酸酯上的任何化学衍生化。术语“衍生物”还包括含有非天然存在的核苷酸和核苷酸类似物的核酸。优选地,核酸的衍生化提高其稳定性。

[0100] “来源于核酸序列的核酸序列”是指其所来源的核酸的变体的核酸。优选地,当关于特定序列的变体的序列代替RNA分子中的特定序列时,其保留了RNA稳定性和/或翻译效率。

[0101] “nt”是核苷酸,或核酸分子中的核苷酸,优选连续核苷酸的缩写。

[0102] 根据本发明,术语“密码子”是指编码核酸中的碱基三联体,其在核糖体的蛋白质合成期间指示接下来将添加哪个氨基酸。

[0103] 术语“转录”及其变化形式涉及这样的过程,在该过程中由RNA聚合酶读取具有特定核酸序列的核酸分子(“核酸模板”),从而使RNA聚合酶产生单链RNA分子。在转录过程中,核酸模板中的遗传信息被转录。核酸模板可以是DNA;但是,例如在从甲病毒核酸模板转录的情况下,模板通常是RNA。随后,可将转录的RNA翻译成蛋白质。根据本发明,术语“转录”包括“体外转录”,其中术语“体外转录”涉及其中在无细胞系统中体外合成RNA,特别是mRNA的过程。优选地,将克隆载体应用于转录物的产生。这些克隆载体通常称为转录载体,并且根据本发明被术语“载体”所涵盖。克隆载体优选是质粒。根据本发明,RNA优选是体外转录的RNA (IVT-RNA),并且可通过合适的DNA模板的体外转录获得。用于控制转录的启动子可以是任何RNA聚合酶的任何启动子。用于体外转录的DNA模板可通过克隆核酸,特别是cDNA,并且将其引入用于体外转录的合适载体中来获得。cDNA可通过RNA的逆转录获得。

[0104] 在转录过程中产生的单链核酸分子通常具有作为模板的互补序列的核酸序列。

[0105] 根据本发明,术语“模板”或“核酸模板”或“模板核酸”通常是指可被复制或转录的核酸序列。

[0106] “从核酸序列转录的核酸序列”和类似术语是指作为模板核酸序列的转录产物的核酸序列,在适当情况下作为完整RNA分子的一部分。通常而言,转录的核酸序列是单链RNA分子。

[0107] 根据本发明,“核酸的3'末端”是指具有游离羟基的末端。在双链核酸(特别是DNA)的示意图中,3'末端始终在右侧。根据本发明,“核酸的5'末端”是指具有游离磷酸基团的末端。在双链核酸(特别是DNA)的示意图中,5'末端始终在左侧。

[0108] 5'末端5' --P-NNNNNN-OH-3' 3'末端3'-HO-NNNNNN-P--5'

[0109] “上游”描述了核酸分子的第一元件相对于该核酸分子的第二元件的相对定位,其中两个元件包含在同一核酸分子中,并且其中第一元件比核酸分子的第二元件更靠近核酸分子的5'末端。第二元件然后被称是在该核酸分子的第一元件的“下游”。可将位于第二元件“上游”的元件同义地称为位于第二元件的“5'”。对于双链核酸分子,关于(+)链给出了如

“上游”和“下游”的指示。

[0110] 根据本发明,“功能性连接”或“功能性连接的”涉及处于功能关系内的连接。如果核酸与另一核酸序列功能相关,则该核酸是“功能性连接的”。例如,如果启动子影响所述编码序列的转录,则启动子与编码序列功能性连接。功能性连接的核酸通常彼此相邻,在适当情况下被另外的核酸序列隔开,并且在特定的实施方案中,通过RNA聚合酶转录得到单个RNA分子(共同转录物)。

[0111] 在一些特定的实施方案中,根据本发明将核酸与表达控制序列功能性连接,所述表达控制序列相对于该核酸可以是同源的或异源的。

[0112] 根据本发明,术语“表达控制序列”包括启动子、核糖体结合序列和其他控制基因的转录或来源RNA的翻译的控制元件。在本发明的特定实施方案中,可调节表达控制序列。表达控制序列的精确结构可根据物种或细胞类型而变化,但是通常包括分别参与起始转录和翻译的5'-非转录序列以及5'-和3'-非翻译序列。更具体地,5'-非转录表达控制序列包括启动子区,所述启动子区包含用于功能性连接的基因的转录控制的启动子序列。表达控制序列还可包括增强子序列或上游激活子序列。DNA分子的表达控制序列通常包括5'-非转录序列以及5'-和3'-非翻译序列,例如TATA盒、加帽序列、CAAT序列等。甲病毒RNA的表达控制序列可包括亚基因组启动子和/或一个或多个保守序列元件。如本文中所述,根据本发明的特定表达控制序列是甲病毒的亚基因组启动子。

[0113] 本文中指定的核酸序列(特别是可转录和编码核酸序列)可与任何表达控制序列(特别是启动子)组合,所述表达控制序列对于所述核酸序列可以是同源的或异源的,术语“同源”是指核酸序列也与表达控制序列是天然功能性连接的事实,以及术语“异源”是指核酸序列并非与表达控制序列是天然功能性连接的事实。

[0114] 如果可转录核酸序列(特别是编码肽或蛋白质的核酸序列)与表达控制序列以这样的方式彼此共价连接,使得可转录核酸序列并且特别是编码核酸序列的转录或表达处于表达控制序列的控制之下或影响之下,则其彼此“功能性”连接。如果要将核酸序列翻译成功能性肽或蛋白质,则诱导表达控制序列功能性连接至编码序列导致所述编码序列的转录,而不导致编码序列中的移码或编码序列不能翻译成所期望的肽或蛋白质。

[0115] 术语“启动子”或“启动子区”是指通过提供对于RNA聚合酶的识别和结合位点来控制转录物(例如包含编码序列的转录物)的合成的核酸序列。启动子区可包含针对参与所述基因转录调节的另外的因子的另外的识别或结合位点。启动子可控制原核或真核基因的转录。启动子可以是“诱导型”并响应于诱导物而启动转录,或者如果转录不受诱导物控制,则可以是“组成型”。如果不存在诱导物,则诱导型启动子仅在很小的程度上表达或完全不表达。在诱导物的存在下,基因被“打开”或转录水平升高。这通常通过特定转录因子的结合来介导。根据本发明的特定启动子是如本文中所述的例如甲病毒的亚基因组启动子。另一些特定启动子是例如甲病毒的基因组正链或负链启动子。

[0116] 术语“核心启动子”是指启动子所包含的核酸序列。核心启动子通常是正确启动转录所需的启动子的最小部分。核心启动子通常包含转录起始位点和RNA聚合酶的结合位点。

[0117] “聚合酶”通常是指能够催化由单体结构单元合成聚合物分子的分子实体。“RNA聚合酶”是能够催化由核糖核苷酸结构单元合成RNA分子的分子实体。“DNA聚合酶”是能够催化由脱氧核糖核苷酸结构单元合成DNA分子的分子实体。对于DNA聚合酶和RNA聚合酶的情

况,分子实体通常是蛋白质或多个蛋白质的组装体或复合物。通常而言,DNA聚合酶基于模板核酸合成DNA分子,所述模板核酸通常是DNA分子。通常而言,RNA聚合酶基于模板核酸合成RNA分子,所述模板分子是DNA分子(在这种情况下,RNA聚合酶是DNA依赖性RNA聚合酶,DdRP),或者是RNA分子(在这种情况下,RNA聚合酶是RNA依赖性RNA聚合酶,RdRP)。

[0118] “RNA依赖性RNA聚合酶”或“RdRP”或“复制酶”是催化从RNA模板的RNA转录的酶。在甲病毒RNA依赖性RNA聚合酶的情况下,基因组RNA的(-)链互补链和(+)链基因组RNA的顺序合成导致RNA复制。因此,RNA依赖性RNA聚合酶被同义地称为“RNA复制酶”或简单地“复制酶”。实际上,RNA依赖性RNA聚合酶通常由除逆转录病毒以外的所有RNA病毒编码。编码RNA依赖性RNA聚合酶的病毒的典型代表是甲病毒。

[0119] 根据本发明,“RNA复制”通常是指基于给定RNA分子(模板RNA分子)的核苷酸序列合成的RNA分子。合成的RNA分子可例如与模板RNA分子相同或互补。通常而言,RNA复制可通过DNA中间体的合成发生,或者可通过RNA依赖性RNA聚合酶(RdRP)介导的RNA依赖性RNA复制直接发生。在甲病毒的情况下,RNA复制不是通过DNA中间体发生的,而是由RNA依赖性RNA聚合酶(RdRP)介导的:模板RNA链(第一RNA链)或其一部分充当模板,用于合成与第一RNA链或其一部分互补的第二RNA链。第二RNA链或其一部分可继而任选地充当模板,用于合成与第二RNA链或其一部分互补的第三RNA链。因此,第三RNA链与第一RNA链或其一部分相同。因此,RNA依赖性RNA聚合酶能够直接合成模板的互补RNA链,并且能够间接合成相同的RNA链(通过互补中间链)。

[0120] 根据本发明,术语“模板RNA”是指可被RNA依赖性RNA聚合酶转录或复制的RNA。

[0121] 根据本发明,术语“基因”是指负责产生一种或更多种细胞产物和/或实现一种或更多种细胞间或细胞内功能的特定核酸序列。更具体地,所述术语涉及核酸部分(通常为DNA;但是在RNA病毒的情况下为RNA),其包含编码特定蛋白质或功能性或结构性RNA分子的核酸。

[0122] 如本文中使用的,“分离的分子”旨在是指基本上不含其他分子(例如其他细胞物质)的分子。根据本发明,术语“分离的核酸”意指核酸已经(i)在体外扩增,例如通过聚合酶链式反应(polymerase chain reaction,PCR),(ii)通过克隆重组产生,(iii)纯化,例如通过切割和凝胶电泳分级,或(iv)合成,例如通过化学合成。分离的核酸是可用于通过重组技术进行操作的核酸。

[0123] 术语“载体”在本文中以其最一般的含义使用,并且包括核酸的任何中间载剂,其例如能够将所述核酸引入原核和/或真核宿主细胞中,并且在适当的情况下被整合到基因组中。这样的载体优选在细胞中复制和/或表达。载体包括质粒、噬菌粒、病毒基因组及其级分。

[0124] 在本发明的上下文中,术语“重组”意指“通过遗传改造制备”。优选地,在本发明的上下文中,“重组对象”(例如重组细胞)不是天然存在的。

[0125] 如本文中使用的术语“天然存在”是指对象可在自然界中存在的事实。例如,存在于生物体(包括病毒)中且可从天然来源中分离并且未在实验室中故意人为修饰的肽或核酸是天然存在的。术语“在自然界中存在”意指“存在于自然界中”,并且包括已知对象以及尚未从自然界中发现和/或分离但将来可能会从天然来源中发现和/或分离的对象。

[0126] 根据本发明,术语“表达”以其最一般的含义使用,并且包括产生RNA和/或蛋白质。

其还包括核酸的部分表达。此外,表达可以是瞬时的或稳定的。对于RNA,术语“表达”或“翻译”涉及细胞核糖体中的过程,通过该过程,编码RNA(例如信使RNA)链指导氨基酸序列的组装,以产生肽或蛋白。

[0127] 根据本发明,术语“mRNA”意指“信使RNA”,并且涉及通常通过使用DNA模板产生并编码肽或蛋白质的转录物。通常而言,mRNA包含5'-UTR、蛋白质编码区、3'-UTR和poly(A)序列。mRNA可从DNA模板通过体外转录而产生。体外转录方法是技术人员已知的。例如,存在许多可商购的体外转录试剂盒。根据本发明,可通过稳定修饰和加帽来对mRNA进行修饰。

[0128] 根据本发明,术语“poly(A)序列”或“poly(A)尾”是指通常位于RNA分子的3'末端的腺苷酸残基的不间断的或间断的序列。不间断的序列的特征是连续的腺苷酸残基。实际上,不间断的poly(A)序列是典型的。尽管poly(A)序列在真核DNA中通常不编码,而是在细胞核中真核转录期间在转录之后通过不依赖于模板的RNA聚合酶连接至RNA的游离3'末端,但是本发明涵盖了由DNA编码的poly(A)序列。

[0129] 根据本发明,关于核酸分子的术语“一级结构”是指核苷酸单体的线性序列。

[0130] 根据本发明,关于核酸分子的术语“二级结构”是指核酸分子的二维表示,其反映了碱基配对;例如在单链RNA分子的情况下,特别是分子内碱基配对。尽管每个RNA分子具有仅一条多核苷酸链,但分子通常以(分子内)碱基对区域为特征。根据本发明,术语“二级结构”包括结构基序,其包括但不限于碱基对、茎、茎环、凸起、环(例如内部环和多分支环)。核酸分子的二级结构可用显示碱基配对的二维图(平面图)表示(关于RNA分子二级结构的更多详细信息,参见Auber et al.,2006;J.Graph Algorithms Appl.10:329-351)。如本文中所述,某些RNA分子的二级结构在本发明的上下文中是相关的。

[0131] 根据本发明,通过使用用于RNA二级结构预测的网络服务器(<http://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html>)进行预测来确定核酸分子,特别是单链RNA分子的二级结构。优选地,根据本发明,关于核酸分子的“二级结构”特别是指通过所述预测确定的二级结构。也可使用MFOLD结构预测(<http://unafold.rna.albany.edu/?q=mfold>)执行或确认该预测。

[0132] 根据本发明,“碱基对”是二级结构的结构基序,其中两个核苷酸碱基通过碱基上的供体和受体位点之间的氢键彼此缔合。互补碱基A:U和G:C通过碱基上供体和受体位点之间的氢键形成稳定的碱基对;A:U和G:C碱基对称为Watson-Crick碱基对。通过碱基G和U(G:U)形成较弱的碱基对(称为Wobble碱基对)。碱基对A:U和G:C称为规范碱基对。其他碱基对如G:U(在RNA中经常出现)和其他稀有碱基对(例如A:C;U:U)称为非规范碱基对。

[0133] 根据本发明,“核苷酸配对”是指彼此缔合以使其碱基形成碱基对(规范或非规范碱基对,优选规范碱基对,最优选Watson-Crick碱基对)的两个核苷酸。

[0134] 根据本发明,关于核酸分子的术语“茎环”或“发夹”或“发夹环”全部可互换地是指核酸分子(通常是单链核酸分子,例如单链RNA)的特定二级结构。由茎环代表的特定二级结构由包含茎和(末端)环(也称为发夹环)的连续核酸序列组成,其中茎由两个相邻的完全或部分互补的序列元件形成,所述序列元件被形成茎环结构的环的短序列(例如3至10个核苷酸)隔开。两个相邻的完全或部分互补的序列可定义为例如茎环元件茎1和茎2。当这两个相邻的完全或部分反向互补的序列(例如茎环元件茎1和茎2)彼此形成碱基对时,形成茎环,这导致由位于茎环元件茎1和茎2之间的短序列在其末端结尾处形成包含未配对环的双链

核酸序列。因此,茎环包含两个茎(茎1和茎2),其在核酸分子的二级结构水平上彼此形成碱基对,并且在核酸分子的一级结构水平上被不是茎1或茎2的一部分的短序列隔开。为了说明,茎环的二维表示类似于棒棒糖形结构(lollipop-shaped structure)。茎环结构的形成需要自身可折回以形成配对双链的序列的存在;配对双链由茎1和茎2形成。配对茎环元件的稳定性通常由长度,能够与茎2的核苷酸形成碱基对(优选规范碱基对,更优选Watson-Crick碱基对)的茎1的核苷酸数目相比于不能与茎2的核苷酸形成这样的碱基对(错配或凸起)的茎1的核苷酸数目来确定。根据本发明,最佳环长度为3至10个核苷酸,更优选4至7个核苷酸,例如4个核苷酸、5个核苷酸、6个核苷酸或7个核苷酸。如果给定核酸序列以茎环为特征,则相应互补核酸序列通常也以茎环为特征。茎环通常由单链RNA分子形成。例如,甲病毒基因组RNA的5'复制识别序列中存在数个茎环。

[0135] 根据本发明,关于核酸分子的特定二级结构(例如茎环)的“破坏”或其变化形式意指特定二级结构不存在或被改变。通常而言,二级结构可由于二级结构的一部分的至少一个核苷酸的改变而被破坏。例如,可通过改变形成茎的一个或更多个核苷酸以使得不能进行核苷酸配对来破坏茎环。

[0136] 根据本发明,“补偿二级结构破坏”或其变化形式是指核酸序列中的一个或更多个核苷酸变化;更通常是指核酸序列中的一个或更多个第二核苷酸变化,所述核酸序列还包含一个或更多个第一核苷酸变化,其特征在在于:尽管在不存在一个或更多个第二核苷酸变化的情况下一个或更多个第一核苷酸变化造成核酸二级结构破坏,但是一个或更多个第一核苷酸变化和一个或更多个第二核苷酸变化的共存不造成核酸二级结构破坏。共存意指一个或更多个第一核苷酸变化和一个或更多个第二核苷酸变化二者的存在。通常而言,一个或更多个第一核苷酸变化和一个或更多个第二核苷酸变化一起存在于同一核酸分子中。在一个具体的实施方案中,补偿二级结构破坏的一个或更多个核苷酸变化是补偿一个或更多个核苷酸配对破坏的一个或更多个核苷酸变化。因此,在一个实施方案中,“补偿二级结构破坏”意指“补偿核苷酸配对破坏”,即一个或更多个核苷酸配对破坏,例如一个或更多个茎环内的一个或更多个核苷酸配对破坏。可通过去除至少一个起始密码子来引入一个或更多个一个或更多个核苷酸配对破坏。补偿二级结构破坏的一个或更多个核苷酸变化中的每一个都是这样的核苷酸变化,其可各自独立地选自一个或更多个核苷酸的缺失、添加、替换和/或插入。在一个示例性实例中,当核苷酸配对A:U被A至C的替换破坏(C和U通常不适合形成核苷酸对)时,则补偿核苷酸配对破坏的核苷酸变化可以是U替换为G,从而能够形成C:G核苷酸配对。因此将U替换为G补偿了核苷酸配对破坏。在一个替代实例中,当核苷酸配对A:U被A至C的替换破坏时,则补偿核苷酸配对破坏的核苷酸变化可以是C替换为A,从而恢复原始A:U核苷酸配对的形成。一般而言,在本发明中,补偿二级结构破坏的那些核苷酸变化优选既不恢复原始核酸序列,也不产生新的AUG三联体。在以上实例集中,相对于C至A的替换,U至G的替换是优选的。

[0137] 根据本发明,关于核酸分子的术语“三级结构”是指由原子坐标所限定的核酸分子的三维结构。

[0138] 根据本发明,核酸(例如RNA,例如rRNA)可编码肽或蛋白质。因此,可转录的核酸序列或其转录物可包含编码肽或蛋白质的开放阅读框(ORF)。

[0139] 根据本发明,术语“编码肽或蛋白质的核酸”意指这样的核酸,如果存在于合适的

环境中,优选存在于细胞内,其可在翻译过程期间指导氨基酸的组装以产生肽或蛋白质。优选地,根据本发明的编码RNA能够与细胞翻译机器相互作用,从而允许编码RNA的翻译以产生肽或蛋白质。

[0140] 根据本发明,术语“肽”包括寡肽和多肽,并且是指包含通过肽键彼此连接的两个或更多个、优选3个或更多个、优选4个或更多个、优选6个或更多个、优选8个或更多个、优选10个或更多个、优选为13个或更多个、优选16个或更多个、优选20个或更多个、并且多至优选50个、优选100个或优选150个连续氨基酸的物质。术语“蛋白质”是指大的肽,优选具有至少151个氨基酸的肽,但是术语“肽”和“蛋白质”在本文中通常用作同义词。

[0141] 根据本发明,术语“肽”和“蛋白质”包括不仅包含氨基酸组分而且还包含非氨基酸组分(例如糖和磷酸结构)的物质,并且还包括包含键(例如酯、硫醚或二硫键)的物质。

[0142] 根据本发明,术语“起始密码子(initiation codon)”和“起始密码子(start codon)”同义地指RNA分子的密码子(碱基三联体),其是由核糖体翻译的潜在第一密码子。这样的密码子通常在真核生物中编码氨基酸甲硫氨酸,并且在原核生物中编码经修饰的甲硫氨酸。真核生物和原核生物中最常见的起始密码子是AUG。除非在本文中特别指出是指除AUG以外的起始密码子,否则关于RNA分子的术语“起始密码子”和“起始密码子”是指密码子AUG。根据本发明,术语“起始密码子”和“起始密码子”也用于指脱氧核糖核酸的相应碱基三联体,即编码RNA的起始密码子的碱基三联体。如果信使RNA的起始密码子是AUG,则编码AUG的碱基三联体是ATG。根据本发明,术语“起始密码子”和“起始密码子”优选是指功能性起始密码子,即,指被核糖体用作或将用作启动翻译的密码子的起始密码子。RNA分子中可能存在未被核糖体用作开始翻译的密码子的AUG密码子,例如由于密码子与帽的距离很短。术语功能性起始密码子不包括这些密码子。

[0143] 根据本发明,术语“开放阅读框的起始密码子(start codon of the open reading frame)”和“开放阅读框的起始密码子(initiation codon of the open reading frame)”是指在编码序列(例如自然界中存在的核酸分子的编码序列)中用作蛋白质合成的起始密码子的碱基三联体。在RNA分子中,开放阅读框的起始密码子通常前面是5' 非翻译区(5' -UTR),尽管这不是严格要求。

[0144] 根据本发明,术语“开放阅读框的天然起始密码子(native start codon of the open reading frame)”和“开放阅读框的天然起始密码子(native initiation codon of the open reading frame)”是指在天然编码序列中用作蛋白质合成的起始密码子的碱基三联体。天然编码序列可以是例如自然界中存在的核酸分子的编码序列。在一些实施方案中,本发明提供了自然界中存在的核酸分子的变体,其特征在于天然起始密码子(存在于天然编码序列中)已被去除(从而其不存在于变体核酸分子中)。

[0145] 根据本发明,“第一AUG”意指信使RNA分子的最上游的AUG碱基三联体,优选被核糖体用作或将要被核糖体用作开始翻译的密码子的信使RNA分子最上游的AUG碱基三联体。相应地,“第一ATG”是指编码第一AUG的编码DNA序列的ATG碱基三联体。在一些情况下,mRNA分子的第一AUG是开放阅读框的起始密码子,即在核糖体蛋白质合成过程中用作起始密码子的密码子。

[0146] 根据本发明,关于核酸变体的某元件的术语“包括去除…”或“以去除…为特征”和类似术语意指与参考核酸分子相比,在核酸变体中所述某元件不是功能性的或不存在。不

受限制地,去除可由某元件的全部或一部分的缺失、某元件的全部或一部分的替换、或某元件的功能或结构性质的改变组成。核酸序列的功能元件的去除要求包含去除的核酸变体在该位置处不表现功能。例如,以去除某起始密码子为特征的RNA变体要求以去除为特征的RNA变体在该位置处不启动核糖体蛋白质合成。核酸序列的结构元件的去除要求包含去除的核酸变体在该位置处不存在该结构元件。例如,以去除某AUG碱基三联体(即,某位置处的AUG碱基三联体)为特征的RNA变体的特征可在于,例如该某AUG碱基三联体的一部分或全部的缺失(例如, Δ AUG),或者该某AUG碱基三联体的一个或更多个核苷酸(A、U、G)被任意一个或更多个不同的核苷酸替换,使得所得变体的核苷酸序列不包含所述AUG碱基三联体。一个核苷酸的合适替换是将AUG碱基三联体转化为GUG、CUG或UUG碱基三联体,或转化为AAG、ACG或AGG碱基三联体,或转化为AUA、AUC或AUU碱基三联体的那些。可相应地选择更多核苷酸的合适替换。

[0147] 根据本发明,术语“自复制病毒”包括能够在宿主细胞中自主复制的RNA病毒。自复制病毒可具有单链RNA(single-stranded RNA, ssRNA)基因组,并包括甲病毒、黄病毒、麻疹病毒(measles virus, MV)和弹状病毒(Rhabdovirus)。甲病毒和黄病毒具有正极性基因组,而麻疹病毒(MV)和弹状病毒的基因组是负链ssRNA。通常而言,自复制病毒是具有在感染细胞之后直接翻译的(+)链RNA基因组的病毒,并且这种翻译提供RNA依赖性RNA聚合酶,其然后产生来自感染的RNA的反义和有义转录物二者。在下文中,通过参考甲病毒来源的载体作为自复制病毒来源的载体的一个实例来举例说明本发明。然而,应理解的是,本发明不限于甲病毒来源的载体。

[0148] 根据本发明,术语“甲病毒”应被广义地理解,并且包括具有甲病毒特征的任何病毒颗粒。甲病毒的特征包括存在(+)链RNA,其编码适合在宿主细胞中复制的遗传信息,包括RNA聚合酶活性。在Strauss&Strauss, 1994, Microbiol. Rev. 58:491-562中描述了许多甲病毒的另外的特征。术语“甲病毒”包括自然界中存在的甲病毒及其任何变体或衍生物。在一些实施方案中,变体或衍生物在自然界中不存在。

[0149] 在一个实施方案中,甲病毒是自然界中存在的甲病毒。通常而言,自然界中存在的甲病毒可感染任一种或更多种真核生物,例如动物(包括脊椎动物(例如人)和节肢动物(例如昆虫))。自然界中存在的甲病毒优选地选自以下:巴马森林病毒复合体(包含巴马森林病毒);东方马脑炎复合体(包含七种抗原类型的东方马脑炎病毒);米德尔堡病毒复合体(包含米德尔堡病毒(Middelburg virus));恩杜姆病毒复合体(包含恩杜姆病毒(Ndumu virus));塞姆利基森林病毒复合体(包含贝巴鲁病毒(Bebaru virus)、基孔肯亚病毒、马雅罗病毒(Mayaro virus)及其亚型乌纳病毒(Una virus)、阿良良病毒(O' Nyong Nyong virus)及其亚型伊博-奥拉病毒(Igbo-Ora virus)、罗斯河病毒(Ross River virus)及其亚型比巴鲁病毒(Bebaru virus)、盖塔病毒(Getah virus)、鹭山病毒(Sagiyama virus)、塞姆利基森林病毒及其亚型Me Tri病毒);委内瑞拉马脑炎复合体(包含卡巴斯欧病毒(Cabassou virus)、沼泽地病毒(Everglades virus)、莫斯达斯佩德拉斯病毒(Mosso das Pedras virus)、穆坎布病毒(Mucambo virus)、帕拉马纳病毒(Paramana virus)、那皮舒纳病毒(Pixuna virus)、里奥内格罗病毒(Rio Negro virus)、特罗卡拉病毒(Trocara virus)及其亚型比朱布里奇病毒(Bijou Bridge virus)、委内瑞拉马脑炎病毒);西方马脑炎复合体(包含奥拉病毒(Aura virus)、巴班肯病毒(Babanki virus)、孜拉加奇病毒

(Kyzylagach virus)、辛德毕斯病毒、奥克尔布病毒 (Ockelbo virus)、沃达罗河病毒 (Whataroa virus)、博吉河病毒 (Buggy Creek virus)、摩根堡病毒 (Fort Morgan virus)、高地J病毒 (Highlands J virus)、西方马脑炎病毒 (Western equine encephalitis virus)); 以及一些未分类的病毒, 包括鲑鱼胰腺疾病病毒 (Salmon pancreatic disease virus); 睡眠疾病病毒 (Sleeping Disease virus); 南部象海豹病毒 (Southern elephant seal virus); 图那特病毒 (Tonate virus)。更优选地, 甲病毒选自以下: 塞姆利基森林病毒复合体 (包含如上所述的病毒类型, 包括塞姆利基森林病毒)、西方马脑炎复合体 (包含如上所述的病毒类型, 包括辛德毕斯病毒)、东方马脑炎病毒 (包含如上所述的病毒类型)、委内瑞拉马脑炎复合体 (包含如上所述的病毒类型, 包括委内瑞拉马脑炎病毒)。

[0150] 在另一个优选实施方案中, 甲病毒是塞姆利基森林病毒。在另一个优选实施方案中, 甲病毒是辛德毕斯病毒。在另一个替代的优选实施方案中, 甲病毒是委内瑞拉马脑炎病毒。

[0151] 在本发明的一些实施方案中, 甲病毒不是自然界中存在的甲病毒。通常而言, 非自然界中存在的甲病毒是自然界中存在的甲病毒的变体或衍生物, 其通过核苷酸序列 (即基因组RNA) 中的至少一个突变而区别于自然界中存在的甲病毒。与自然界中存在的甲病毒相比, 核苷酸序列中的突变可选自一个或更多个核苷酸的插入、替换或缺失。核苷酸序列中的突变可与由核苷酸序列编码的多肽或蛋白质中的突变相关或不相关。例如, 非自然界中存在的甲病毒可以是减毒甲病毒。非自然界中存在的减毒甲病毒是这样的甲病毒, 其通常在其核苷酸序列中具有至少一个突变, 通过该突变, 其区别于自然界中存在的甲病毒, 并且其没有完全没有感染性, 或者具有感染性但具有较低疾病产生能力或完全没有疾病产生能力。作为说明性实例, TC83是区别于自然界中存在的委内瑞拉马脑炎病毒 (VEEV) 的减毒甲病毒 (McKinney et al., 1963, Am. J. Trop. Med. Hyg. 12: 597-603)。

[0152] 甲病毒属的成员也可根据其在人中的相对临床特征进行分类: 主要与脑炎相关的甲病毒, 以及主要与发热、皮疹和多关节炎相关的甲病毒。

[0153] 术语“甲病毒的”意指在存在于甲病毒中, 或来源于甲病毒或来源于甲病毒 (例如通过遗传改造)。

[0154] 根据本发明, “SFV”代表塞姆利基森林病毒。根据本发明, “SIN”或“SINV”代表辛德毕斯病毒。根据本发明, “VEE”或“VEEV”代表委内瑞拉马脑炎病毒。

[0155] 根据本发明, 术语“甲病毒的”是指来源于甲病毒的实体。为了说明, 甲病毒的蛋白质可指存在于甲病毒中的蛋白质和/或由甲病毒编码的蛋白质; 并且甲病毒的核酸序列可指存在于甲病毒中的核酸序列和/或由甲病毒编码的核酸序列。优选地, “甲病毒的”核酸序列是指“甲病毒的基因组的”和/或“甲病毒的基因组RNA的”核酸序列。

[0156] 根据本发明, 术语“甲病毒RNA”是指以下任意一种或更多种: 甲病毒基因组RNA (即 (+) 链)、甲病毒基因组RNA的互补链 (即 (-) 链)、和亚基因组转录物 (即 (+) 链) 或其任何片段。

[0157] 根据本发明, “甲病毒基因组”是指甲病毒的基因组 (+) 链RNA。

[0158] 根据本发明, 术语“天然甲病毒序列”和类似术语通常是指天然存在的甲病毒 (自然界中存在的甲病毒) 的 (例如核酸) 序列。在一些实施方案中, 术语“天然甲病毒序列”也包括减毒甲病毒的序列。

[0159] 根据本发明,术语“5' 复制识别序列”优选是指与自复制病毒基因组(例如甲病毒基因组)的5' 片段相同或同源的连续核酸序列,优选核糖核酸序列。“5' 复制识别序列”是可被复制酶(例如甲病毒复制酶)识别的核酸序列。术语5' 复制识别序列包括天然的5' 复制识别序列及其功能等同物,例如,例如自然界中存在的自复制病毒(例如自然界中存在的甲病毒)的5' 复制识别序列的功能变体。根据本发明,功能等同物包括如本文中所述的以除去至少一个起始密码子为特征的5' 复制识别序列的衍生物。5' 复制识别序列是合成甲病毒基因组RNA的(-)链互补链所必需的,并且是基于(-)链模板合成(+)链病毒基因组RNA所必需的。天然5' 复制识别序列通常至少编码nsP1的N端片段;但不包含编码nsP1234的整个开放阅读框。考虑到天然的5' 复制识别序列通常至少编码nsP1的N端片段的事实,天然5' 复制识别序列通常包含至少一个起始密码子,通常为AUG。在一个实施方案中,5' 复制识别序列包含甲病毒基因组的保守序列元件1(conserved sequence element 1,CSE 1)或其变体和甲病毒基因组的保守序列元件2(conserved sequence element 2,CSE 2)或其变体。5' 复制识别序列通常能够形成四个茎环(stem loop,SL),即SL1、SL2、SL3、SL4。这些茎环的编号从5' 复制识别序列的5' 末端开始。

[0160] 术语“保守序列元件”或“CSE”是指存在于甲病毒RNA中的核苷酸序列。这些序列元件称为“保守的”,因为在不同甲病毒的基因组中存在直系同源物,并且不同甲病毒的直系同源CSE优选具有高百分比的序列同一性和/或相似的二级或三级结构。术语CSE包括CSE 1、CSE 2、CSE 3和CSE 4。

[0161] 根据本发明,术语“CSE 1”或“44-nt CSE”同义地是指从(-)链模板合成(+)链所需的核苷酸序列。术语“CSE 1”是指(+)链上的序列;并且CSE 1的互补序列(在(-)链上)用作(+)链合成的启动子。优选地,术语CSE 1包括甲病毒基因组的最5' 核苷酸。CSE 1通常形成保守的茎环结构。不希望受特定理论的束缚,认为对于CSE 1,二级结构比一级结构(即线性序列)更重要。在模型甲病毒辛德毕斯病毒的基因组RNA中,CSE 1由44个核苷酸的连续序列组成,该序列由基因组RNA的最5' 的44个核苷酸形成(Strauss&Strauss,1994, Microbiol.Rev.58:491-562)。

[0162] 根据本发明,术语“CSE 2”或“51-nt CSE”同义地是指从(+)链模板合成(-)链所需的核苷酸序列。(+)链模板通常是甲病毒基因组RNA或RNA复制子(应注意,不包含CSE 2的亚基因组RNA转录物不用作(-)链合成的模板)。在甲病毒基因组RNA中,CSE 2通常位于nsP1的编码序列内。在模型甲病毒辛德毕斯病毒的基因组RNA中,51-nt CSE位于基因组RNA的核苷酸位置155至205(Frolov et al.,2001,RNA,vol.7,pp.1638-1651)。CSE 2通常形成两个保守的茎环结构。这些茎环结构被命名为茎环3(SL3)和茎环4(SL4),因为其分别是甲病毒基因组RNA的5' 末端开始计数的甲病毒基因组RNA的第三和第四保守茎环。不希望受特定理论的束缚,认为对于CSE 2,二级结构比一级结构(即线性序列)更重要。

[0163] 根据本发明,术语“CSE 3”或“连接序列”同义地指来源于甲病毒基因组RNA并包含亚基因组RNA的起始位点的核苷酸序列。(-)链中该序列的互补序列起促进亚基因组RNA转录的作用。在甲病毒基因组RNA中,CSE 3通常与编码nsP4的C端片段的区域重叠,并延伸到位于编码结构蛋白的开放阅读框上游的短非编码区。

[0164] 根据本发明,术语“CSE 4”或“19-nt保守序列”或“19-nt CSE”同义地指来自甲病毒基因组RNA的核苷酸序列,其在甲病毒基因组的3' 非翻译区中的poly(A)序列的立即上

游。CSE 4通常由19个连续核苷酸组成。不希望受特定理论的束缚,CSE 4被理解为充当启动(-)链合成的核心启动子(José et al., 2009, Future Microbiol. 4:837-856);和/或甲病毒基因组RNA的CSE 4和poly (A) 尾被理解为共同作用以有效进行(-)链合成(Hardy&Rice, 2005, J. Virol. 79:4630-4639)。

[0165] 根据本发明,术语“亚基因组启动子”或“SGP”是指核酸序列(例如编码序列)上游(5')的核酸序列,其通过提供对于RNA聚合酶,通常是RNA依赖性RNA聚合酶,特别是功能性甲病毒非结构蛋白的识别和结构位点来控制所述核酸序列的转录。SGP可包含对于另外的因子的另外的识别或结合位点。亚基因组启动子通常是正链RNA病毒(例如甲病毒)的遗传元件。甲病毒的亚基因组启动子是病毒基因组RNA中包含的核酸序列。亚基因组启动子的一般特征在于,其允许在RNA依赖性RNA聚合酶(例如功能性甲病毒非结构蛋白)的存在下开始转录(RNA合成)。RNA(-)链(即甲病毒基因组RNA的互补链)充当用于合成(+)链亚基因组转录物的模板,并且(+)链亚基因组转录物的合成通常在亚基因组启动子处或附近开始。如本文中使用的术语“亚基因组启动子”不限于在包含这样的亚基因组启动子的核酸中的任何特定定位。在一些实施方案中,SGP与CSE 3相同或与CSE 3重叠或包含CSE 3。

[0166] 术语“亚基因组转录物”或“亚基因组RNA”同义地是指由于使用RNA分子作为模板(“模板RNA”)的转录而可获得的RNA分子,其中模板RNA包含控制亚基因组转录物的转录的亚基因组启动子。亚基因组转录物可在存在RNA依赖性RNA聚合酶,特别是功能性甲病毒非结构蛋白的情况下获得。例如,术语“亚基因组转录物”可以是指甲病毒感染的细胞中使用甲病毒基因组RNA的(-)链互补链作为模板制备的RNA转录物。然而,如本文中使用的术语“亚基因组转录物”不限于此,并且还包括通过使用异源RNA作为模板可获得的转录物。例如,也可通过使用根据本发明的含有SGP的复制子的(-)链互补链作为模板来获得亚基因组转录物。因此,术语“亚基因组转录物”可以是指通过转录甲病毒基因组RNA的片段可获得的RNA分子,以及通过转录根据本发明的复制子的片段可获得的RNA分子。

[0167] 术语“自体”用于描述来源于相同对象的任何事物。例如,“自体细胞”是指来源于相同对象的细胞。将自体细胞引入对象是有利的,因为这些细胞克服了免疫学屏障,否则会导致排斥。

[0168] 术语“同种异体”用于描述来源于相同物种的不同个体的任何事物。当一个或多个基因座上的基因不不同时,两个或多个个体被认为是同种异体。

[0169] 术语“同基因的”用于描述来源于具有相同基因型的个体或组织(即相同双胞胎或相同近交系的动物,或者其组织或细胞)的任何事物。

[0170] 术语“异源”用于描述由多个不同元件组成的事物。作为一个实例,将一个个体的细胞引入不同个体就构成了异源移植。异源基因是来源于除对象以外的基因。

[0171] 在用于鉴定序列变化的方法中可使用的细胞是任何合适的细胞,其中具有或不具有任何核苷酸修饰的rRNA可被复制和/或翻译。细胞可以是哺乳动物细胞,例如人细胞。细胞可组成型表达复制酶,所述复制酶识别存在于rRNA中的用于复制的序列,或者可瞬时表达这样的复制酶。

[0172] 以下提供了本发明的各个特征的具体和/或优选的变型。本发明还考虑了通过将对于本发明的两种或更多种特征描述的两种或更多种具体和/或优选变体进行组合产生的那些实施方案作为一些特别优选实施方案。

[0173] RNA复制子

[0174] 可复制RNA (rRNA) 是能够由RNA依赖性RNA聚合酶(复制酶)复制的RNA,其凭借包含可由复制酶识别的核苷酸序列使得RNA被复制。rRNA不一定编码复制酶,使得rRNA可被顺式复制(通过编码的复制酶)或者反式复制(通过以另一种方式提供的复制酶,例如编码单独的复制酶的核酸)。术语“RNA复制子”、“复制子”和“可复制的RNA分子”可互换使用。

[0175] 在一个实施方案中,可复制RNA (rRNA) 分子包含自复制的单链正义病毒的经修饰调节区,与参考的经修饰调节区相比包含序列变化,所述序列变化恢复或提高包含至少一个经修饰核苷酸的rRNA分子的功能。这些变化可通过本文中所述的用于鉴定这样的序列变化的方法来鉴定。在一个实施方案中,经修饰调节区是甲病毒调节区,例如5' 或3' 调节区。在一个实施方案中,甲病毒5' 调节区的序列变化包含对应于SEQ ID NO:2的第67、244、245、246、248位的一个或更多个位置处的点突变。在一个实施方案中,5' 调节区是VEEV甲病毒的5' 调节区。在一个实施方案中,5' 调节区具有SEQ ID NO:1或2中所述的序列。在一个实施方案中,5' 调节区包含具有SEQ ID NO:2中所述序列的VEEV 5' 调节区的第67、244、245、246、248位中的一个或更多个处的点突变或者对应于SEQ ID NO:2中那些位置的另一种甲病毒的5' 调节区中位置的点突变。在一个实施方案中,rRNA分子在5' 调节区中还包含在SEQ ID NO:2的第4位处的点突变。

[0176] 在多个实施方案中,点突变可位于SEQ ID NO:2的第67位,第67和244位,第67和246位,第67和248位,第67、245和248位,第4和67位,第4、67和244位,第4、67和246位,第4、67和248位,第4、67、245和248位处,或者其任意组合处,或者位于对应于SEQ ID NO:2中这样的位置的甲病毒中的位置处。

[0177] 在多个实施方案中,点突变可位于SEQ ID NO:2的以下位置中一个或更多个处或者位置的组合处,或者在甲病毒的5' 调节区中相应的位置处:67,244,245,246,248,67和244,67和245,67和246,67和248,244和245,244和246,244和248,245和246,245和248,246和248,67、244和245,67、244和246,67、244和248,67、245和246,67、245和248,67、246和248,244、245和246,244、245和248,244、246和248,245、246和248,67、244、245和246,67、244、245和248,67、244、246和248,67、245、246和248,244、245、246和248,67、244、245、246和248。在每个前述位置或位置的组合中,点突变还可位于SEQ ID NO:2的第4位处或位于甲病毒的5' 调节区中相应的位置处。

[0178] 在一个实施方案中,位于多个位置处的特定点突变可以是G4A、A67C、G244A、C245A、G246A或C248A。

[0179] 在一个实施方案中,RNA复制子可包含内部核糖体进入位点(IRES)和编码来自自复制病毒的功能性非结构蛋白的开放阅读框,其中IRES控制功能性非结构蛋白例如复制酶的表达。优选地,RNA复制子包含允许由功能性非结构蛋白复制的序列元件。在一个实施方案中,自复制病毒是甲病毒,并且允许由功能性非结构蛋白复制的序列元件来源于甲病毒。

[0180] 甲病毒复制酶具有加帽酶功能,并且通常而言,基因组以及亚基因组(+)链RNA被加帽。5' 帽用于保护mRNA免于降解,并将核糖体亚基以及细胞因子导向mRNA,以在mRNA上形成核糖核蛋白复合体,然后可从附近的起始密码子开始翻译。该复杂的过程在文献中有广泛的描述(Jackson et al.,2010,Nat Rev Mol Biol;Vol 10:113-127)。尽管帽依赖性翻译的机制非常复杂和有效,但细胞具有完全或部分独立于5' 帽的启动翻译的方式

(Thompson 2012; Trends in Microbiology 20:558-566)。因此,在导致帽依赖性翻译的整体下调的细胞应激的情况下,通常在IRES的帮助下,细胞仍然可优先表达选定的基因。

[0181] 病毒还进化出不同的方式以利用细胞机制来翻译病毒基因。由于病毒感染通常由细胞感知,导致细胞抗病毒应答(干扰素应答;应激应答),因此许多病毒也利用帽非依赖性翻译,特别是RNA病毒。帽非依赖性翻译确保了细胞应激应答时病毒RNA翻译的优势,使病毒有机会完成其生命周期并从受感染的细胞中释放出来。

[0182] 内部核糖体进入位点(IRES)是形成合适二级结构的RNA序列,其在翻译起始密码子(AUG或其他)附近吸引预起始复合物。文献中描述了四类具有共同特征的IRES。原型IRES(prototypic IRES)是脊髓灰质炎病毒IRES(I型)、脑心肌炎病毒(encephalomyocarditis virus,EMCV)IRES(II型)、丙型肝炎病毒(hepatitis C virus,HCV)IRES(III型)和存在于双顺反子病毒(dicistroviruses)的基因间区中的IRES(IV型)(Thompson,2012;Trends in Microbiology 20:558-566;Lozano et al.,2018;Open Biology 8:180155)。

[0183] I型至III型IRES的共同之处在于其在AUG起始密码子处起始翻译,而IV型IRES在非AUG密码子(例如GCU)处起始。因此,I型至III型需要起始物tRNA,其在eIF2/GTP的帮助下递送甲硫氨酸(eIF2/GTP/Met-tRNA^{Met})。在应激的情况下,eIF2激酶的激活使eIF2的 α 亚基磷酸化,抑制在AUG处起始的翻译。因此,由IV型IRES指导的翻译不受eIF2磷酸化的抑制。

[0184] 术语“内部核糖体进入位点”,缩写为“IRES”,涉及RNA元件,其将核糖体募集至mRNA的内部区域以帽非依赖性方式开始翻译。IRES通常位于RNA病毒的5'-UTR中。然而,来自双顺反子病毒科(dicistroviridae)的病毒的mRNA具有两个开放阅读框(ORF),并且各自的翻译由两个不同的IRES指导。还表明了一些哺乳动物细胞mRNA也具有IRES。这些细胞IRES元件被认为位于编码基因的真核mRNA,所述基因参与应激存活和其他对存活至关重要的过程。IRES元件的位置通常在5'-UTR中,但也可在mRNA的其他位置出现。

[0185] 术语“内部核糖体进入位点”包括存在于微小RNA病毒科(Picornaviridae)的病毒中的IRES,所述病毒例如脊髓灰质炎病毒(poliovirus,PV)和脑心肌炎病毒以及致病性病毒,包括人免疫缺陷病毒、丙型肝炎病毒(HCV)和口蹄疫病毒。尽管这些病毒IRES包含不同的序列,但其中的许多具有类似的二级结构并通过类似的机制开始翻译。另外,IRES的活性通常需要称为IRES反式作用因子(IRES-transacting factor,ITAF)的另一些因子的帮助。基于翻译起始因子(initiation factor,IF)和ITAF的结构和需要,病毒IRES分为四种类型,如本文中所述。任何这些IRES类型根据本发明都是可用的,其中IV型IRES是特别优选的。

[0186] 两组病毒IRES,I型和II型,不能直接与40S小核糖体亚基结合。相反地,其通过不同的ITAF募集40S小核糖体亚基,并在帽依赖性翻译中需要典型的IF(即eIF2、eIF3、eIF4A、eIF4B和eIF4G)。I型和II型IRES之间的主要区别在于需要40S核糖体扫描,其中40S核糖体扫描对于II型IRES是不必要的。I型IRES的一些实例包括存在于脊髓灰质炎病毒(PV)和鼻病毒中的IRES。II型IRES的一些实例包括存在于脑心肌炎病毒(EMCV)、口蹄疫病毒(foot-and-mouth disease virus,FMDV)和Theiler鼠脑脊髓炎病毒(Theiler's murine encephalomyelitis virus,TMEV)中的IRES。

[0187] III型IRES可直接与具有特化RNA结构的40S小核糖体亚基相互作用,但其活性通常需要数种IF的协助,包括eIF2和eIF3以及起始物Met-tRNAⁱ。一些实例包括存在于丙型肝炎

炎病毒(HCV)、经典猪瘟病毒(classical swine fever virus,CSFV)和猪捷申病毒(porcine teschovirus,PTV)中的IRES。

[0188] IV型病毒IRES通常具有强活性,并且可在没有另外的ITAF或甚至eIF2/Met-tRNAi/GTP三元复合体的情况下从非AUG起始密码子起始翻译。这些IRES被折叠成紧凑的结构,直接与40S小核糖体亚基相互作用。一些实例包括存在于双顺反子病毒例如蟋蟀麻痹病毒(cricket paralysis virus,CrPV)、大豆尺蠖病毒(plautia stali intestine virus,PSIV)和桃拉综合征病毒(Taura-Syndrom-Virus,TSV)中的IRES。

[0189] 术语“内部核糖体进入位点”,还包括存在于细胞mRNA中的IRES,其中的许多编码在应激应答中例如在凋亡、有丝分裂、缺氧和营养限制的状况中所需的蛋白质。基于核糖体募集的机制,细胞IRES可大致分为两种类型:I型IRES通过结合在顺式元件上的ITAF例如RNA结合基序和N-6-甲基腺苷(N-6-methyladenosine,m6A)修饰与核糖体相互作用,而II型IRES包含与18S rRNA配对以募集核糖体的短顺式元件。

[0190] 在一个实施方案中,本文中所述的rRNA可具有经修饰核苷酸/核苷/骨架修饰。如本文中使用的术语“RNA修饰”可以是指化学修饰,其包括骨架修饰以及糖修饰或碱基修饰。

[0191] 在本上下文中,如本文中限定的经修饰的rRNA分子可包含核苷酸类似物/修饰,例如骨架修饰、糖修饰或碱基修饰。与本发明相关的骨架修饰是这样的修饰,其中如本文中限定的rRNA分子中包含的核苷酸骨架的磷酸酯被化学修饰。与本发明相关的糖修饰是如本文中限定的rRNA分子核苷酸糖的化学修饰。此外,与本发明相关的碱基修饰是rRNA分子核苷酸碱基部分的化学修饰。在本上下文中,核苷酸类似物或修饰优选选自适用于转录和/或翻译的核苷酸类似物。

[0192] 糖修饰:经修饰核苷和核苷酸可在糖部分中修饰,所述经修饰核苷和核苷酸可并入到如本文中所述的经修饰的rRNA分子中。例如,2' 羟基(OH)可用许多不同的“氧基”或“脱氧”取代基修饰或替换。“氧基”-2' 羟基修饰的一些实例包括但不限于烷氧基或芳氧基(-OR,例如R=H、烷基、环烷基、芳基、芳烷基、杂芳基或糖);聚乙二醇(polyethyleneglycol,PEG)、-O(CH₂CH₂O)_nCH₂CH₂OR;“锁”核酸(“locked”nucleic acid,LNA),其中2' 羟基例如通过亚甲基桥连接至同一核糖糖的4' 碳上;以及氨基(-O-氨基,其中氨基例如NRR可以是烷基氨基、二烷基氨基、杂环基、芳基氨基、二芳基氨基、杂芳基氨基或二杂芳基氨基、乙二胺、聚氨基)或者氨基烷氧基。“脱氧”修饰包括氢、氨基(例如NH₂;烷基氨基、二烷基氨基、杂环基、芳基氨基、二芳基氨基、杂芳基氨基、二杂芳基氨基或氨基酸);或者氨基可通过接头连接至糖,其中接头包含一个或更多个原子C、N和O。糖基还可包含一个或更多个碳,所述碳具有与核糖中相应的碳相反的立体化学构型。因此,经修饰的RNA分子可包含含有例如阿拉伯糖作为糖的核苷酸。

[0193] 骨架修饰:在经修饰核苷和核苷酸中磷酸骨架可进一步被修饰,其可并入到如本文中所述的经修饰的RNA分子中。骨架的磷酸基团可通过用不同取代基替换一个或更多个氧原子来修饰。另外,经修饰核苷和核苷酸可包括用如本文中所述的经修饰的磷酸酯完全替换未经修饰的磷酸酯部分。经修饰的磷酸基团的一些实例包括但不限于硫代磷酸酯、硒代磷酸酯、硼烷磷酸酯(borano phosphate)、硼代磷酸酯(borano phosphate ester)、膦酸氢酯、氨基磷酸酯、烷基或芳基膦酸酯和磷酸三酯。二硫代磷酸酯具有两个由硫替换的非连接氧。磷酸酯接头也可通过用氮(桥连的氨基磷酸酯)、硫(桥连的硫代磷酸酯)和碳(桥连的

亚甲基磷酸酯) 替换连接氧来修饰。

[0194] 碱基修饰: 经修饰核苷和核苷酸还可在核碱基部分中修饰, 所述经修饰核苷和核苷酸可并入到如本文中所述的经修饰的rRNA分子中。在RNA中存在的核碱基的一些实例包括但不限于腺嘌呤、鸟嘌呤、胞嘧啶和尿嘧啶。例如, 本文中所述的核苷和核苷酸可在大沟(major groove)表面上进行化学修饰。在一些实施方案中, 大沟化学修饰可包括氨基、硫醇基、烷基或卤素基团。

[0195] 在本发明的特定实施方案中, 核苷酸类似物/修饰选自碱基修饰, 其优选地选自2-氨基-6-氯嘌呤核苷-5'-三磷酸、2-氨基嘌呤-核苷-5'-三磷酸; 2-氨基腺苷-5'-三磷酸、2'-氨基-2'-脱氧-胞苷-三磷酸、2-硫代胞苷-5'-三磷酸、2-硫代尿苷-5'-三磷酸、2'-氟代胸苷-5'-三磷酸、2'-0-甲基肌苷-5'-三磷酸、4-硫代-尿苷-5'-三磷酸、5-氨基烯丙基胞苷-5'-三磷酸、5-氨基烯丙基尿苷-5'-三磷酸、5-溴代胞苷-5'-三磷酸、5-溴代尿苷-5'-三磷酸、5-溴代-2'-脱氧胞苷-5'-三磷酸、5-溴代-2'-脱氧尿苷-5'-三磷酸、5-碘代胞苷-5'-三磷酸、5-碘代-2'-脱氧胞苷-5'-三磷酸、5-碘代尿苷-5'-三磷酸、5-碘代-2'-脱氧尿苷-5'-三磷酸、5-甲基胞苷-5'-三磷酸、5-甲基尿苷-5'-三磷酸、5-丙炔基-2'-脱氧胞苷-5'-三磷酸、5-丙炔基-2'-脱氧尿苷-5'-三磷酸、6-氮杂胞苷-5'-三磷酸、6-氮杂尿苷-5'-三磷酸、6-氯嘌呤核苷-5'-三磷酸、7-脱氮杂-腺苷-5'-三磷酸、7-脱氮杂鸟苷-5'-三磷酸、8-氮杂腺苷-5'-三磷酸、8-叠氮基腺苷-5'-三磷酸、苯并咪唑-核苷-5'-三磷酸、N1-甲基腺苷-5'-三磷酸、N1-甲基鸟苷-5'-三磷酸、N6-甲基腺苷-5'-三磷酸、O6-甲基鸟苷-5'-三磷酸、假尿苷-5'-三磷酸、或嘌呤霉素-5'-三磷酸、黄嘌呤核苷-5'-三磷酸。用于碱基修饰的特别优选考虑的核苷酸选自由以下组成的经碱基修饰的核苷酸的组: 5-甲基胞苷-5'-三磷酸、7-脱氮杂鸟苷-5'-三磷酸、5-溴代胞苷-5'-三磷酸、和假尿苷-5'-三磷酸。在一些实施方案中, 经修饰核苷包括吡啶-4-酮核糖核苷、5-氮杂-尿苷、2-硫代-5-氮杂-尿苷、2-硫代尿苷、4-硫代-假尿苷、2-硫代-假尿苷、5-羟基尿苷、3-甲基尿苷、5-羧甲基-尿苷、1-羧甲基-假尿苷、5-丙炔基-尿苷、1-丙炔基-假尿苷、5-牛磺酸甲基尿苷、1-牛磺酸甲基-假尿苷、5-牛磺酸甲基-2-硫代尿苷、1-牛磺酸甲基-4-硫代-尿苷、5-甲基-尿苷、1-甲基-假尿苷、4-硫代-1-甲基-假尿苷、2-硫代-1-甲基-假尿苷、1-甲基-1-脱氮杂-假尿苷、2-硫代-1-甲基-1-脱氮杂-假尿苷、二氢尿苷、二氢-假尿苷、2-硫代-二氢尿苷、2-硫代-二氢假尿苷、2-甲氧基-尿苷、2-甲氧基-4-硫代-尿苷、4-甲氧基-假尿苷、和4-甲氧基-2-硫代-假尿苷。

[0196] 在一些实施方案中, 经修饰核苷包括5-氮杂-胞苷、假异胞苷、3-甲基-胞苷、N4-乙酰基胞苷、5-甲酰基胞苷、N4-甲基胞苷、5-羟甲基胞苷、1-甲基-假异胞苷、吡咯并-胞苷、吡咯并-假异胞苷、2-硫代-胞苷、2-硫代-5-甲基-胞苷、4-硫代-假异胞苷、4-硫代-1-甲基-假异胞苷、4-硫代-1-甲基-1-脱氮杂-假异胞苷、1-甲基-1-脱氮杂-假异胞苷、泽布拉林(zebularine)、5-氮杂-泽布拉林、5-甲基-泽布拉林、5-氮杂-2-硫代-泽布拉林、2-硫代-泽布拉林、2-甲氧基-胞苷、2-甲氧基-5-甲基-胞苷、4-甲氧基-假异胞苷、和4-甲氧基-1-甲基-假异胞苷。

[0197] 在另一些实施方案中, 经修饰核苷包括2-氨基嘌呤、2,6-二氨基嘌呤、7-脱氮杂-腺嘌呤、7-脱氮杂-8-氮杂-腺嘌呤、7-脱氮杂-2-氨基嘌呤、7-脱氮杂-8-氮杂-2-氨基嘌呤、7-脱氮杂-2,6-二氨基嘌呤、7-脱氮杂-8-氮杂-2,6-二氨基-嘌呤、1-甲基腺苷、N6-甲基腺苷、N6-异戊烯基腺苷、N6-(顺式-羟基异戊烯基)腺苷、2-甲硫基-N6-(顺式-羟基异戊烯基)

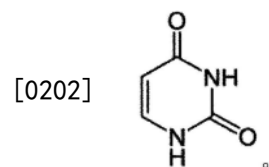
腺苷、N6-甘氨酸基(glycinyI)氨基甲酰基腺苷、N6-苏氨酸基氨基甲酰基腺苷、2-甲基-硫代-N6-苏氨酸基氨基甲酰基腺苷、N6,N6-二甲基腺苷、7-甲基腺嘌呤、2-甲硫基-腺嘌呤、和2-甲氧基-腺嘌呤。在另一些实施方案中,经修饰核苷包括肌苷、1-甲基-肌苷、怀俄苷(wyosine)、怀丁苷(wybutosine)、7-脱氮杂-鸟苷、7-脱氮杂-8-氮杂-鸟苷、6-硫代-鸟苷、6-硫代-7-脱氮杂-鸟苷、6-硫代-7-脱氮杂-8-氮杂-鸟苷、7-甲基-鸟苷、6-硫代-7-甲基-鸟苷、7-甲基肌苷、6-甲氧基-鸟苷、1-甲基鸟苷、N2-甲基鸟苷、N2,N2-二甲基鸟苷、8-氧代-鸟苷、7-甲基-8-氧代-鸟苷、1-甲基-6-硫代-鸟苷、N2-甲基-6-硫代-鸟苷、和N2,N2-二甲基-6-硫代-鸟苷。

[0198] 在一些实施方案中,核苷酸可在大沟表面上修饰,并且可包括用甲基或卤素基团替换尿嘧啶的C-5上的氢。在具体的实施方案中,经修饰核苷是5'-0-(1-硫代磷酸酯)-腺苷、5'-0-(1-硫代磷酸酯)-胞苷、5'-0-(1-硫代磷酸酯)-鸟苷、5'-0-(1-硫代磷酸酯)-尿苷或5'-0-(1-硫代磷酸酯)-假尿苷。

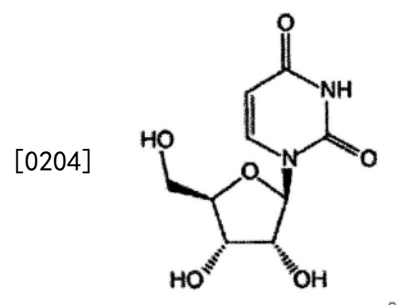
[0199] 在另一些实施方案中,经修饰的rRNA可包含选自以下的核苷修饰:6-氮杂-胞苷、2-硫代-胞苷、 α -硫代-胞苷、假异胞苷、5-氨基烯丙基-尿苷、5-碘代-尿苷、N1-甲基-假尿苷、5,6-二氢尿苷、 α -硫代-尿苷、4-硫代-尿苷、6-氮杂-尿苷、5-羟基-尿苷、脱氧-胸苷、5-甲基-尿苷、吡咯并-胞苷、肌苷、 α -硫代-鸟苷、6-甲基-鸟苷、5-甲基-胞苷(cytidine)、8-氧代-鸟苷、7-脱氮杂-鸟苷、N1-甲基-腺苷、2-氨基-6-氯代-嘌呤、N6-甲基-2-氨基-嘌呤、假异胞苷、6-氯代-嘌呤、N6-甲基-腺苷、 α -硫代-腺苷、8-叠氮基-腺苷、7-脱氮杂-腺苷。

[0200] 在某些优选实施方案中,rRNA包含替换至少一个(例如每个)尿苷的经修饰核苷。

[0201] 如本文中使用的术语“尿嘧啶”描述了在RNA核酸中可存在的核碱基中的一种。尿嘧啶的结构是:

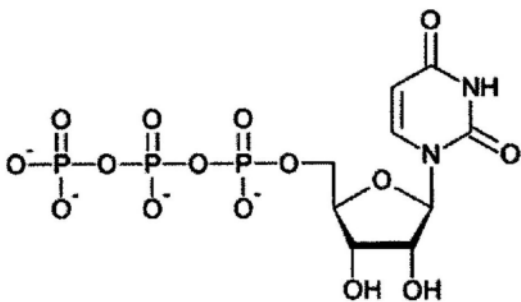


[0203] 如本文中使用的术语“尿苷”描述了在RNA中可存在的核苷中的一种。尿苷的结构是:



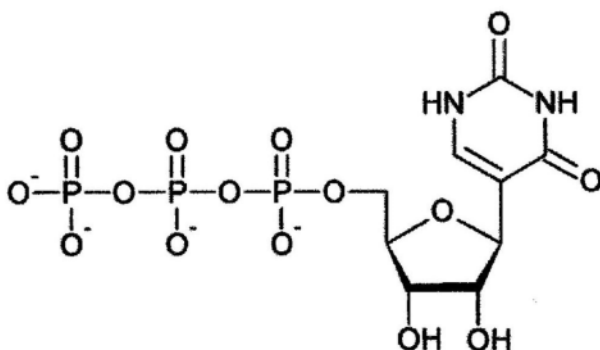
[0205] UTP(尿苷5'-三磷酸)具有以下结构:

[0206]



[0207] 假UTP (假尿苷5' -三磷酸) 具有以下结构:

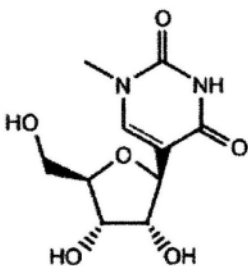
[0208]



[0209] “假尿苷”是经修饰核苷的一个实例,其为尿苷的异构体,其中尿嘧啶通过碳-碳键而不是氮-碳糖苷键与戊糖环连接。

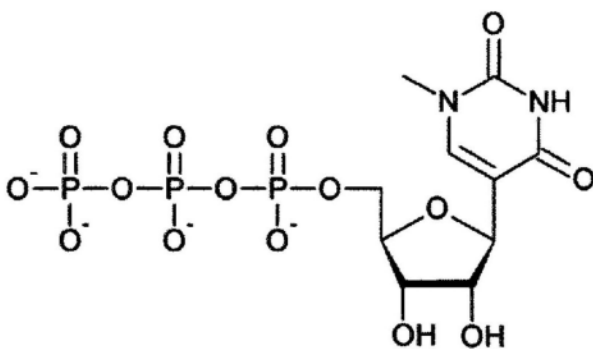
[0210] 另一种示例性的经修饰核苷是N1-甲基-假尿苷 (m1ψ),其具有以下结构:

[0211]

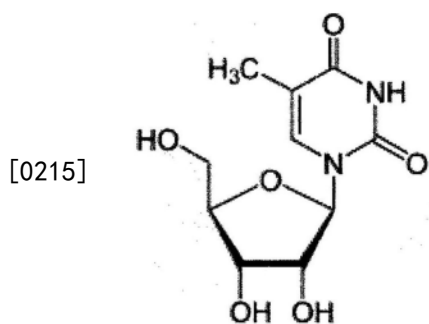


[0212] N1-甲基-假UTP具有以下结构:

[0213]



[0214] 另一种示例性的经修饰核苷是5-甲基-尿苷 (m5U),其具有以下结构:



[0216] 在某些优选实施方案中,本文中所述rRNA中的一个或多个尿苷由经修饰的核苷替换。在一些实施方案中,经修饰的核苷是经修饰的尿苷。

[0217] 在某些优选实施方案中,RNA包含替换了至少一个尿苷的经修饰核苷。在一些实施方案中,RNA包含替换了每个尿苷的经修饰核苷。

[0218] 在某些优选的实施方案中,经修饰的核苷独立地选自假尿苷(ψ)、N1-甲基-假尿苷($m1\psi$)、和5-甲基-尿苷($m5U$)。在一些实施方案中,经修饰的核苷包含假尿苷(ψ)。在一些实施方案中,经修饰的核苷包含N1-甲基-假尿苷($m1\psi$)。在一些实施方案中,经修饰核苷包含5-甲基-尿苷($m5U$)。在一些实施方案中,RNA可包含多于一种类型的经修饰核苷,并且经修饰核苷独立地选自假尿苷(ψ)、N1-甲基-假尿苷($m1\psi$)、和5-甲基-尿苷($m5U$)。在一些实施方案中,经修饰核苷包含假尿苷(ψ)和N1-甲基-假尿苷($m1\psi$)。在一些实施方案中,经修饰核苷包含假尿苷(ψ)和5-甲基-尿苷($m5U$)。在一些实施方案中,经修饰核苷包含N1-甲基-假尿苷($m1\psi$)和5-甲基-尿苷($m5U$)。在一些实施方案中,经修饰核苷包含假尿苷(ψ)、N1-甲基-假尿苷($m1\psi$)、和5-甲基-尿苷($m5U$)。

[0219] 在某些优选的实施方案中,替换rRNA中一个或多个(例如全部)尿苷的经修饰核苷可以是以下的一种或更多种:3-甲基-尿苷(m^3U)、5-甲氧基-尿苷(mo^5U)、5-氮杂-尿苷、6-氮杂-尿苷、2-硫代-5-氮杂-尿苷、2-硫代-尿苷(s^2U)、4-硫代-尿苷(s^4U)、4-硫代-假尿苷、2-硫代-假尿苷、5-羟基-尿苷(ho^5U)、5-氨基烯丙基-尿苷、5-卤代-尿苷(例如5-碘代-尿苷或5-溴代-尿苷)、尿苷5-羟乙酸(cmo^5U)、尿苷5-羟乙酸甲酯($mcmo^5U$)、5-羧甲基-尿苷(cm^5U)、1-羧甲基-假尿苷、5-羧基羟甲基-尿苷(chm^5U)、5-羧基羟甲基-尿苷甲酯($mchm^5U$)、5-甲氧基羰基甲基-尿苷(mcm^5U)、5-甲氧基羰基甲基-2-硫代-尿苷(mcm^5s^2U)、5-氨基甲基-2-硫代-尿苷(nm^5s^2U)、5-甲基氨基甲基-尿苷(mnm^5U)、1-乙基-假尿苷、5-甲基氨基甲基-2-硫代-尿苷(mnm^5s^2U)、5-甲基氨基甲基-2-硒代-尿苷(mnm^5se^2U)、5-氨基甲酰基甲基-尿苷(ncm^5U)、5-羧甲基氨基甲基-尿苷($cmnm^5U$)、5-羧甲基氨基甲基-2-硫代-尿苷($cmnm^5s^2U$)、5-丙炔基-尿苷、1-丙炔基-假尿苷、5-牛磺酸甲基-尿苷(τm^5U)、1-牛磺酸甲基-假尿苷、5-牛磺酸甲基-2-硫代-尿苷(τm^5s^2U)、1-牛磺酸甲基-4-硫代-假尿苷、5-甲基-2-硫代-尿苷(m^5s^2U)、1-甲基-4-硫代-假尿苷($m^1s^4\psi$)、4-硫代-1-甲基-假尿苷、3-甲基-假尿苷($m^3\psi$)、2-硫代-1-甲基-假尿苷、1-甲基-1-脱氮杂-假尿苷、2-硫代-1-甲基-1-脱氮杂-假尿苷、二氢尿苷(D)、二氢假尿苷、5,6-二氢尿苷、5-甲基-二氢尿苷(m^5D)、2-硫代-二氢尿苷、2-硫代-二氢假尿苷、2-甲氧基-尿苷、2-甲氧基-4-硫代-尿苷、4-甲氧基-假尿苷、4-甲氧基-2-硫代-假尿苷、N1-甲基-假尿苷、3-(3-氨基-3-羧基丙基)尿苷(acp^3U)、1-甲基-3-(3-氨基-3-羧基丙基)假尿苷($acp^3\psi$)、5-(异戊烯基氨基甲基)尿苷(inm^5U)、5-(异戊烯基氨基甲基)-2-硫代-尿苷(inm^5s^2U)、 α -硫代-尿苷、2'-O-甲基-尿苷(Um)、5,2'-O-二甲基-尿苷(m^5Um)、2'-

0-甲基-假尿苷(ψ m)、2-硫代-2'-0-甲基-尿苷(s^2 Um)、5-甲氧基羰基甲基-2'-0-甲基-尿苷(mcm^5 Um)、5-氨基甲酰基甲基-2'-0-甲基尿苷(ncm^5 Um)、5-羧甲基氨基甲基-2'-0-甲基-尿苷($cmnm^5$ Um)、3,2'-0-二甲基-尿苷(m^3 Um)、5-(异戊烯基氨基甲基)-2'-0-甲基-尿苷(inm^5 Um)、1-硫代-尿苷、脱氧胸苷、2'-F-阿糖-尿苷、2'-F-尿苷、2'-OH-阿糖-尿苷、5-(2-甲酯基乙烯基)尿苷、5-[3-(1-E-丙烯基氨基)尿苷、或本领域已知的任何其他经修饰尿苷。

[0220] 在一个实施方案中,rRNA包含其他经修饰核苷或包含另一些经修饰核苷,例如如上所述的那些经修饰胞苷。例如,在一个实施方案中,在rRNA中5-甲基胞苷由胞苷部分或完全替换,优选完全替换。在一个实施方案中,rRNA包含5-甲基胞苷以及选自假尿苷(ψ)、N1-甲基-假尿苷($m1\psi$)、和5-甲基-尿苷($m5U$)中的一种或更多种。在一个实施方案中,rRNA包含5-甲基胞苷和N1-甲基-假尿苷($m1\psi$)。在一些实施方案中,rRNA包含替换了每个胞苷的5-甲基胞苷以及替换了每个尿苷的N1-甲基-假尿苷($m1\psi$)。

[0221] 功能性非结构蛋白

[0222] 术语“非结构蛋白”涉及由病毒编码但不属于病毒颗粒的蛋白质。该术语通常包括病毒用于自身复制的多种酶和转录因子,例如RNA复制酶或另一些模板导向的聚合酶。术语“非结构蛋白”包括每种和每个共翻译或翻译后修饰形式,包括非结构蛋白的碳水化合物修饰(例如糖基化)和脂质修饰形式,并且优选涉及“甲病毒非结构蛋白”。

[0223] 在一些实施方案中,术语“甲病毒非结构蛋白”是指甲病毒来源的任意一种或更多种单独的非结构蛋白($nsP1$ 、 $nsP2$ 、 $nsP3$ 、 $nsP4$),或者是指多蛋白,其包含甲病毒来源的多于一种非结构蛋白的多肽序列。在一些实施方案中,“甲病毒非结构蛋白”是指 $nsP123$ 和/或 $nsP4$ 。在另一些实施方案中,“甲病毒非结构蛋白”是指 $nsP1234$ 。在一个实施方案中,由开放阅读框编码的目的蛋白质由全部 $nsP1$ 、 $nsP2$ 、 $nsP3$ 和 $nsP4$ 组成,作为一个单个的任选可切割的多蛋白: $nsP1234$ 。在一个实施方案中,由开放阅读框编码的目的蛋白质由 $nsP1$ 、 $nsP2$ 和 $nsP3$ 组成,作为一个单个的任选可切割的多蛋白: $nsP123$ 。在该实施方案中, $nsP4$ 可以是另一种目的蛋白质,并且可由另一个开放阅读框编码。

[0224] 在一些实施方案中,非结构蛋白能够例如在宿主细胞中形成复合体或缔合体(association)。在一些实施方案中,“甲病毒非结构蛋白”是指 $nsP123$ (同义 $P123$)和 $nsP4$ 的复合体或缔合体。在一些实施方案中,“甲病毒非结构蛋白”是指 $nsP1$ 、 $nsP2$ 和 $nsP3$ 的复合体或缔合体。在一些实施方案中,“甲病毒非结构蛋白”是指 $nsP1$ 、 $nsP2$ 、 $nsP3$ 和 $nsP4$ 的复合体或缔合体。在一些实施方案中,“甲病毒非结构蛋白”是指选自 $nsP1$ 、 $nsP2$ 、 $nsP3$ 和 $nsP4$ 中的任意一种或更多种的复合体或缔合体。在一些实施方案中,甲病毒非结构蛋白至少包含 $nsP4$ 。

[0225] 术语“复合体”或“缔合体”是指在空间上邻近的两个或更多个相同或不同的蛋白质分子。复合体的蛋白质优选彼此直接或间接物理或理化接触。复合体或缔合体可由多种不同的蛋白质(异源多聚体)和/或一种特定蛋白质的多个拷贝(同源多聚体)组成。在甲病毒非结构蛋白的背景下,术语“复合体或缔合体”描述了许多至少两种蛋白质分子,其中至少一种是甲病毒非结构蛋白。复合体或缔合体可由一种特定蛋白质(同源多聚体)和/或多种不同蛋白质(异源多聚体)的多个拷贝组成。在多聚体的情况下,“多”意指多于一个,例如两个、三个、四个、五个、六个、七个、八个、九个、十个或多于十个。

[0226] 术语“功能性非结构蛋白”包括具有复制酶功能的非结构蛋白。因此,“功能性非结构蛋白”包括甲病毒复制酶。“复制酶功能”包括RNA依赖性RNA聚合酶(RdRP)的功能,即能够基于(+)链RNA模板催化(-)链RNA的合成的酶和/或能够基于(-)链RNA模板催化(+)链RNA的合成的酶。因此,术语“功能性非结构蛋白”可以是指使用(+)链(例如基因组)RNA作为模板合成(-)链RNA的蛋白质或复合体、使用基因组RNA的(-)链互补链作为模板合成新的(+)链RNA的蛋白质或复合体和/或使用基因组RNA的(-)链互补链的片段作为模板合成亚基因组转录物的蛋白质或复合体。功能性非结构蛋白可另外具有一种或更多种另外的功能,例如如蛋白酶(用于自切割)、解旋酶、末端腺苷转移酶(用于poly(A)尾添加)、甲基转移酶和鸟苷酸转移酶(guanylyltransferase)(用于提供具有5'帽的核酸)、核定位位点、三磷酸酶(Gould et al.,2010,Antiviral Res.87:111-124;Rupp et al.,2015,J.Gen.Virol.96:2483-500)。

[0227] 术语“复制酶”包括RNA依赖性RNA聚合酶。根据本发明,术语“复制酶”包括“甲病毒复制酶”,包括来自天然存在的甲病毒(存在于自然界中的甲病毒)的RNA依赖性RNA聚合酶和来自甲病毒变体或衍生物(例如来自减毒甲病毒)的RNA依赖性RNA聚合酶。

[0228] 术语“复制酶”包括由甲病毒感染的细胞表达或由已经用编码甲病毒复制酶的核酸转染的细胞表达的甲病毒复制酶的所有变体,特别是翻译后经修饰的变体、构象、同种型和同源物。此外,术语“复制酶”包括已经产生并且可通过重组方法产生的复制酶的所有形式。例如,可通过重组方法产生包含有助于在实验室中检测和/或纯化复制酶的标签的复制酶,所述标签例如myc标签、HA标签或寡组氨酸标签(His标签)。

[0229] 任选地,甲病毒复制酶另外在功能上由结合以下任意一种或更多种的能力限定:甲病毒保守序列元件1(CSE 1)或其互补序列、保守序列元件2(CSE 2)或其互补序列、保守序列元件3(conserved sequence element3,CSE 3)或其互补序列、保守序列元件4(conserved sequence element 4,CSE 4)或其互补序列。优选地,复制酶能够结合CSE 2[即(+)链]和/或CSE 4[即(+)链]或者结合CSE 1的互补链[即(-)链]和/或CSE 3的互补链[即(-)链]。

[0230] 甲病毒复制酶的来源不限于任何特定的甲病毒。在一个优选实施方案中,甲病毒复制酶包含来自塞姆利基森林病毒的非结构蛋白,所述塞姆利基森林病毒包括天然存在的塞姆利基森林病毒和塞姆利基森林病毒的变体或衍生物,例如减毒的塞姆利基森林病毒。在一个替代的优选实施方案中,甲病毒复制酶包含来自辛德毕斯病毒的非结构蛋白,所述辛德毕斯病毒包括天然存在的辛德毕斯病毒和辛德毕斯病毒的变体或衍生物,例如减毒的辛德毕斯病毒。在一个替代的优选实施方案中,甲病毒复制酶包含来自委内瑞拉马脑炎病毒(VEEV)的非结构蛋白,所述委内瑞拉马脑炎病毒包括天然存在的VEEV和VEEV的变体或衍生物,例如减毒的VEEV。在一个替代的优选实施方案中,甲病毒复制酶包含来自基孔肯雅病毒(CHIKV)的非结构蛋白,所述基孔肯雅病毒包括天然存在的CHIKV和CHIKV的变体或衍生物,例如减毒的CHIKV。

[0231] 复制酶还可包含来自多于一种病毒(例如来自多于一种甲病毒)的非结构蛋白。因此,包含甲病毒非结构蛋白并具有复制酶功能的异源复合体或缔合体同样包含在本发明内。仅出于举例说明目的,复制酶可包含来自第一甲病毒的一种或更多种非结构蛋白(例如nsP1、nsP2)和来自第二甲病毒的一种或更多种非结构蛋白(nsP3、nsP4)。来自多于一种的

不同甲病毒的非结构蛋白可由单独的开放阅读框编码或者可由单个开放阅读框编码为多蛋白,例如nsP1234。

[0232] 在一些实施方案中,功能性非结构蛋白能够在其中表达功能性非结构蛋白的细胞中形成膜复制复合体和/或泡。

[0233] 如果由根据本发明的核酸分子编码功能性非结构蛋白,即具有复制酶功能的非结构蛋白,则优选的是复制子的亚基因组启动子(如果存在的话)与所述复制酶相容。在这种情况下相容意味着复制酶能够识别亚基因组启动子(如果存在的话)。在一个实施方案中,当亚基因组启动子对复制酶来源的病毒是天然的,即这些序列的天然来源是相同的病毒时实现了这一点。在一个替代实施方案中,亚基因组启动子对病毒复制酶来源的病毒不是天然的,前提是病毒复制酶能够识别亚基因组启动子。换言之,复制酶与亚基因组启动子相容(交叉病毒相容性)。关于来源于不同甲病毒的亚基因组启动子和复制酶的交叉病毒相容性的实例在本领域中是已知的。亚基因组启动子和复制酶的任何组合都是可能的,只要存在交叉病毒相容性即可。交叉病毒相容性可容易地由从事本发明工作的技术人员在适合于从亚基因组启动子进行RNA合成的条件下通过将待测试的复制酶与RNA一起孵育来测试,其中RNA具有待测试的亚基因组启动子。如果制备亚基因组转录物,则确定亚基因组启动子和复制酶是相容的。已知交叉病毒相容性的多种实例。

[0234] 复制子可优选由功能性非结构蛋白复制。特别地,编码功能性非结构蛋白的RNA复制子可被该复制子编码的功能性非结构蛋白复制。在一个优选实施方案中,RNA复制子包含编码功能性甲病毒非结构蛋白的开放阅读框。在一个实施方案中,复制子包含编码目的蛋白质的另一个开放阅读框。该实施方案特别适用于产生根据本发明的目的蛋白质的一些方法。在一个实施方案中,编码目的蛋白质的另一个开放阅读框位于5'复制识别序列的下游且位于IRES的上游(且位于编码来自自复制病毒的功能性非结构蛋白的开放阅读框的上游)且/或位于编码来自自复制病毒的功能非结构蛋白的开放阅读框的下游。编码位于5'复制识别序列的下游且位于IRES的上游(且位于编码来自自复制病毒的功能性非结构蛋白的开放阅读框的上游)的目的蛋白质的另一个开放阅读框可表达为具有由5'复制识别序列编码的序列的融合蛋白。编码位于5'复制识别序列的下游且位于IRES的上游(且位于编码来自自复制病毒的功能性非结构蛋白的开放阅读框的上游)的目的蛋白质的另一个开放阅读框可由或可不由亚基因组启动子控制。位于编码来自自复制病毒的功能性非结构蛋白的开放阅读框的下游的编码一个或更多个目的蛋白质的一个或更多个另外的开放阅读框通常由亚基因组启动子控制。

[0235] 优选的是编码功能性非结构蛋白的开放阅读框不与5'复制识别序列重叠。在一个实施方案中,编码功能性非结构蛋白的开放阅读框不与亚基因组启动子(如果存在的话)重叠。其实施方案在WO 2017/162460中公开,通过引用并入本文。

[0236] 复制和蛋白质编码区所需的序列元件的解偶联

[0237] 由于编码nsP1234的开放阅读框与甲病毒基因组的5'复制识别序列(nsP1的编码序列)重叠并且通常还与包含CSE 3的亚基因组启动子(nsP4的编码序列)重叠,因此多种甲病毒来源的载体难以开发。

[0238] 本文中所述的RNA复制子通常包含通过复制酶进行复制所需的序列元件,特别是5'复制识别序列。在一个实施方案中,非结构蛋白的编码序列在IRES的控制下,并且因此

IRES位于非结构蛋白的编码序列的上游。因此,在一个实施方案中,通常与甲病毒非结构蛋白的N端片段的编码序列重叠的5'复制识别序列位于IRES的上游,并且与非结构蛋白的编码序列不重叠。

[0239] 在一个实施方案中,5'复制识别序列的编码序列(例如nsP1编码序列)在框内与位于IRES上游的目的基因融合。

[0240] 在一个实施方案中,5'复制识别序列不编码任何蛋白质或其片段,例如甲病毒非结构蛋白或其片段。因此,在根据本发明的RNA复制子中,通过复制酶进行复制所需的序列元件与蛋白质编码区可以不偶联。与天然病毒基因组RNA(例如,天然甲病毒基因组RNA)相比,可通过去除5'复制识别序列中的至少一个起始密码子来实现解偶联。

[0241] 因此,rRNA可包含5'复制识别序列,其中5'复制识别序列的特征在于,与天然病毒5'复制识别序列(例如,天然甲病毒的5'复制识别序列)相比,其包含至少一个起始密码子的去除。

[0242] 5'复制识别序列的特征在于,与天然病毒5'复制识别序列相比,其包含至少一个起始密码子的去除,在本文中可称为“经修饰的5'复制识别序列”或“根据本发明的5'复制识别序列”。如下文所述,根据本发明的5'复制识别序列可任选地通过存在一个或更多个另外的核苷酸变化来表征,例如通过本发明的方法所检测的核苷酸变化。

[0243] 能够被复制酶(优选甲病毒复制酶)复制的核酸构建体称为可复制RNA或复制子。根据本发明,术语“复制子”定义了可通过RNA依赖性RNA聚合酶复制的RNA分子,其在没有DNA中间体的情况下产生RNA复制子的一个或更多个相同或基本相同的拷贝。“没有DNA中间体”意指在形成RNA复制子的拷贝的过程中没有形成复制子的脱氧核糖核酸(DNA)拷贝或互补链,和/或在形成RNA复制子的拷贝或其互补链的过程中没有将脱氧核糖核酸(DNA)分子用作模板。复制酶功能通常由功能性非结构蛋白例如功能性甲病毒非结构蛋白提供。

[0244] 根据本发明,术语“可被复制”和“能够被复制”通常描述了可制备核酸的一个或更多个相同或基本相同的拷贝。当与术语“复制酶”一起使用时,例如在“能够被复制酶复制”中,术语“可被复制”和“能够被复制”描述了核酸分子(例如RNA复制子)关于复制酶的功能特征。这些功能特征包括以下的至少一种:(i)复制酶能够识别复制子和(ii)复制酶能够充当RNA依赖性RNA聚合酶(RdRP)。优选地,复制酶能够(i)识别复制子并且(ii)充当RNA依赖性RNA聚合酶二者。

[0245] 表述“能够识别”描述复制酶能够与复制子物理缔合,并且优选地,复制酶能够与复制子结合(通常为非共价)。术语“结合”可意指复制酶具有与以下任一种或更多种结合的能力:保守序列元件1(CSE 1)或其互补序列(如果被复制子包含)、保守序列元件2(CSE 2)或其互补序列(如果被复制子包含)、保守序列元素3(CSE 3)或其互补序列(如果被复制子包含)、保守序列元素4(CSE 4)或其互补序列(如果被复制子包含)。优选地,复制酶能够与CSE 2[即(+)链]和/或CSE 4[即(+)链]结合,或者与CSE 1的互补序列[即(-)链]和/或CSE 3的互补序列[即(-)链]结合。

[0246] 在一个实施方案中,表述“能够充当RdRP”意指复制酶能够催化病毒基因组(+)链RNA的(-)链互补链的合成,其中(+)链RNA具有模板功能,和/或复制酶能够催化(+)链病毒基因组RNA的合成,其中(-)链RNA具有模板功能。通常而言,表述“能够充当RdRP”还可包括复制酶能够催化(+)链亚基因组转录物的合成,其中(-)链RNA具有模板功能,并且其中(+)

链亚基因组转录物的合成通常在亚基因组启动子处开始。在一个实施方案中,所述病毒是甲病毒。

[0247] 表述“能够结合”和“能够充当RdRP”是指在正常生理条件下的能力。特别地,其是指表达功能性非结构蛋白或已被编码功能性非结构蛋白的核酸转染的细胞内部的条件。细胞优选是真核细胞。结合的能力和/或充当RdRP的能力可通过实验来测试,例如在无细胞的体外系统或真核细胞中。任选地,所述真核细胞是来自以下物种的细胞:对于所述物种,代表复制酶来源的特定病毒具有感染性。例如,当使用来自对人具有感染性的特定病毒的病毒复制酶时,正常生理条件是人细胞中的条件。更优选地,真核细胞(在一个实例中是人细胞)来自代表复制酶来源的特定病毒对其具有感染性的同一组织或器官。

[0248] 根据本发明,“与天然甲病毒序列相比”和类似的术语是指作为天然甲病毒序列的变体的序列。该变体本身通常不是天然甲病毒序列。

[0249] 在一个实施方案中,RNA复制子包含3'复制识别序列。3'复制识别序列是可被功能性非结构蛋白识别的核酸序列。换言之,功能性非结构蛋白能够识别3'复制识别序列。优选地,3'复制识别序列位于复制子的3'末端(如果复制子不包含poly(A)尾),或紧接在poly(A)尾的上游(如果复制子包含poly(A)尾)。在一个实施方案中,3'复制识别序列由CSE 4组成或包含CSE 4。

[0250] 在一个实施方案中,5'复制识别序列和3'复制识别序列能够在功能性非结构蛋白的存在下指导根据本发明的RNA复制子的复制。因此,当单独存在或优选一起存在时,这些识别序列在功能性非结构蛋白的存在下指导RNA复制子的复制。

[0251] 优选地,提供能够识别复制子的5'复制识别序列和3'复制识别序列二者的功能性非结构蛋白。在一个实施方案中,当3'复制识别序列对于功能性甲病毒非结构蛋白来源的甲病毒是天然的并且当5'复制识别序列对于功能性甲病毒非结构蛋白来源的甲病毒是天然的或者是对于功能性甲病毒非结构蛋白来源的甲病毒是天然的5'复制识别序列的变体时,实现了这一点。天然意指这些序列的天然来源是相同的甲病毒。在一个替代实施方案中,5'复制识别序列和/或3'复制识别序列对于功能性甲病毒非结构蛋白来源的甲病毒不是天然的,前提是功能性甲病毒非结构蛋白能够识别复制子的5'复制识别序列和3'复制识别序列二者。换言之,功能性甲病毒非结构蛋白与5'复制识别序列和3'复制识别序列二者相容。当非天然功能性甲病毒非结构蛋白能够识别相应的序列或序列元件时,则该功能性甲病毒非结构蛋白被认为是相容的(交叉病毒相容性)。(3'/5')复制识别序列和CSE分别与功能性甲病毒非结构蛋白的任何组合都是可能的,只要存在交叉病毒相容性即可。通过在适合于RNA复制的条件下,例如在合适的宿主细胞中将待测试的功能性甲病毒非结构蛋白与RNA一起孵育,交叉病毒相容性可容易地由从事本发明工作的技术人员测试,其中所述RNA具有待测试的3'和5'复制识别序列。如果发生复制,则确定(3'/5')复制识别序列与功能性甲病毒非结构蛋白是相容的。

[0252] 去除5'复制识别序列内的至少一个起始密码子提供了数个优点。编码nsP1*(nsP1的N端片段)的核酸序列中起始密码子的缺少通常导致nsP1*不翻译。此外,由于nsP1*不翻译,编码目的蛋白质的开放阅读框(“G0I 2”)是核糖体可接近的最上游的开放阅读框;因此,当复制子存在于细胞中时,翻译在编码目的基因的开放阅读框(RNA)的第一AUG处开始。

[0253] 至少一个起始密码子的去除可通过本领域已知的任何合适的方法来实现。例如,

可在计算机上(in silico)设计编码根据本发明的复制子(即特征在于起始密码子的去除)的合适的DNA分子,并且然后在体外合成(基因合成);或者,可通过编码复制子的DNA序列的定点诱变来获得合适的DNA分子。在任何情况下,相应的DNA分子可用作体外转录的模板,从而提供根据本发明的复制子。

[0254] 与天然5'复制识别序列相比,至少一个起始密码子的去除没有特别限制,并且可选自任何核苷酸修饰,包括一个或多个核苷酸的替换(包括在DNA水平上,起始密码子的A和/或T和/或G的替换);一个或多个核苷酸的缺失(包括在DNA水平上,起始密码子的A和/或T和/或G的缺失),以及一个或多个核苷酸的插入(包括在DNA水平上,在起始密码子的A和T之间和/或T和G之间一个或多个核苷酸的插入)。无论核苷酸修饰是替换、插入还是缺失,核苷酸修饰都不得导致形成新的起始密码子(作为说明性示例:在DNA水平上,插入必须不是ATG的插入)。

[0255] 以至少一个起始密码子的去除为特征的RNA复制子的5'复制识别序列(即,根据本发明的经修饰的5'复制识别序列)优选是自然界中存在的甲病毒的基因组的5'复制识别序列的变体。在一个实施方案中,根据本发明的经修饰的5'复制识别序列优选地特征在于,与自然界中存在的至少一种甲病毒的基因组的5'复制识别序列具有80%或更高、优选85%或更高、更优选90%或更高、甚至更优选95%或更高的序列同一性程度。

[0256] 在一个实施方案中,可以以去除至少一个起始密码子为特征的RNA复制子的5'复制识别序列包含与甲病毒的5'末端(即甲病毒基因组的5'末端)的约250个核苷酸同源的序列。在一个优选实施方案中,其包含与甲病毒的5'末端(即甲病毒基因组的5'末端)的约250至500个,优选约300至500个核苷酸同源的序列。“甲病毒基因组的5'末端”意指从甲病毒基因组的最上游核苷酸处开始并包括其的核酸序列。换言之,甲病毒基因组的最上游核苷酸指定为核苷酸号1,并且例如“甲病毒基因组的5'末端的250个核苷酸”意指甲病毒基因组的核苷酸1至250。在一个实施方案中,RNA复制子的5'复制识别序列的特征在于与自然界中存在的至少一种甲病毒的基因组的5'末端的至少250个核苷酸具有80%或更高、优选85%或更高、更优选90%或更高、甚至更优选95%或更高的序列同一性程度。至少250个核苷酸包括例如250个核苷酸、300个核苷酸、400个核苷酸、500个核苷酸。

[0257] 自然界中存在的甲病毒的5'复制识别序列通常特征在于至少一个起始密码子和/或保守的二级结构基序。例如,塞姆利基森林病毒(SFV)的天然5'复制识别序列包含五个特定的AUG碱基三联体。根据Frolov et al., 2001, RNA7:1638-1651通过MFOLD的分析显示,预计塞姆利基森林病毒的自然5'复制识别序列形成四个茎环(SL),称为茎环1至4(SL1、SL2、SL3、SL4)。根据Frolov等人通过MFOLD的分析显示,预计不同的甲病毒辛德毕斯病毒的自然5'复制识别序列也形成四个茎环:SL1、SL2、SL3、SL4。

[0258] 已知甲病毒基因组的5'末端包含能够通过功能性甲病毒非结构蛋白复制甲病毒基因组的序列元件。在本发明的一个实施方案中,RNA复制子的5'复制识别序列包含与甲病毒的保守序列元件1(CSE 1)同源的序列和/或与甲病毒的保守序列元件2(CSE 2)同源的序列。

[0259] 甲病毒基因组RNA的保守序列元件2(CSE 2)通常由SL3和SL4代表,其前面是至少包含编码甲病毒非结构蛋白nsP1的第一氨基酸残基的天然起始密码子的SL2。然而,在本说明书中,在一些实施方案中,甲病毒基因组RNA的保守序列元件2(CSE 2)是指跨越SL2至SL4

并且包含编码甲病毒非结构蛋白nsP1的第一氨基酸残基的天然起始密码子的区域。在一个优选实施方案中, RNA复制子包含CSE 2或与CSE 2同源的序列。在一个实施方案中, RNA复制子包含与CSE 2同源的序列, 所述序列的优选特征在于与自然界中存在的至少一种甲病毒的CSE 2的序列具有80%或更高、优选85%或更高、更优选90%或更高、甚至更优选95%或更高的序列同一性程度。

[0260] 在一个实施方案中, 5' 复制识别序列包含与甲病毒的CSE 2同源的序列。甲病毒的CSE 2可包含来自甲病毒的非结构蛋白的开放阅读框的片段。

[0261] 因此, 在一个实施方案中, RNA复制子的特征在于其包含与来自甲病毒的非结构蛋白或其片段的开放阅读框同源的序列。与非结构蛋白或其片段的开放阅读框同源的序列通常是自然界中存在的甲病毒的非结构蛋白或其片段的开放阅读框的变体。在一个实施方案中, 与非结构蛋白或其片段的开放阅读框同源的序列优选地特征在于与自然界中存在的至少一种甲病毒的非结构蛋白或其片段的开放阅读框具有80%或更高、优选85%或更高、更优选90%或更高、甚至更优选95%或更高的序列同一性程度。

[0262] 在一个实施方案中, 与本发明的复制子所包含的非结构蛋白的开放阅读框同源的序列不包含非结构蛋白的天然起始密码子, 并且更优选地, 不包含非结构蛋白的任何起始密码子。在一个实施方案中, 与CSE 2同源的序列的特征在于与天然甲病毒CSE 2序列相比去除了所有起始密码子。因此, 与CSE 2同源的序列优选不包含任何起始密码子。

[0263] 当与开放阅读框同源的序列不包含任何起始密码子时, 与开放阅读框同源的序列本身不是开放阅读框, 因为其不能用作翻译模板。

[0264] 在一个实施方案中, 5' 复制识别序列包含与来自甲病毒的非结构蛋白或其片段的开放阅读框同源的序列, 其中与来自甲病毒的非结构蛋白或其片段的开放阅读框同源的序列的特征在于, 与天然甲病毒序列相比, 其包含至少一个起始密码子的去除。

[0265] 在一个实施方案中, 与来自甲病毒的非结构蛋白或其片段的开放阅读框同源的序列的特征在于, 其包含非结构蛋白的开放阅读框的至少天然起始密码子的去除。优选地, 其特征在于, 其包含编码nsP1的开放阅读框的至少天然起始密码子的去除。

[0266] 天然起始密码子是AUG碱基三联体, 当宿主细胞中存在RNA时, 宿主细胞中核糖体上的翻译在此处开始。换言之, 天然起始密码子是例如已经用包含天然起始密码子的RNA接种的宿主细胞中核糖体蛋白质合成过程中翻译的第一碱基三联体。在一个实施方案中, 宿主细胞是来自真核物种的细胞, 所述真核物种是包含天然甲病毒5' 复制识别序列的特定甲病毒的天然宿主。在一个实施方案中, 宿主细胞是来自细胞系“BHK21[C13] (ATCC® CCL10™)”的BHK21细胞, 可从美国弗吉尼亚州马纳萨斯美国典型培养物保藏中心 (American Type Culture Collection, Manassas, Virginia, USA) 获得。

[0267] 许多甲病毒的基因组已经完全测序并且可公开获得, 而且由这些基因组编码的非结构蛋白序列也可公开获得。这样的序列信息允许在计算机上确定天然起始密码子。

[0268] 在一个实施方案中, 与来自甲病毒的非结构蛋白或其片段的开放阅读框同源的序列的特征在于, 其包含非结构蛋白的开放阅读框的天然起始密码子以外的一个或更多个起始密码子的去除。在一个实施方案中, 所述核酸序列另外的特征在于天然起始密码子的去除。例如, 除了去除天然起始密码子之外, 还可去除任意一个或两个或三个或四个或多于四个 (例如五个) 起始密码子。

[0269] 如果复制子的特征在于去除非结构蛋白的开放阅读框的天然起始密码子并且任选地去除的除天然起始密码子以外的一个或更多个起始密码子,则与开放阅读框同源的序列本身不是开放阅读框,因为其不能用作翻译模板。

[0270] 优选地除了去除天然起始密码子之外,去除的除天然起始密码子以外的一个或更多个起始密码子优选地选自具有开始翻译的潜力的AUG碱基三联体。具有开始翻译的潜力的AUG碱基三联体可称为“潜在起始密码子”。给定的AUG碱基三联体是否具有开始翻译的潜力可在计算机上或在基于细胞的体外测定中确定。

[0271] 在一个实施方案中,在计算机上确定给定的AUG碱基三联体是否具有开始翻译的潜力:在该实施方案中,检查核苷酸序列,并且如果AUG碱基三联体是AUGG序列的一部分,优选Kozak序列的一部分,则确定其具有开始翻译的潜力。

[0272] 在一个实施方案中,在基于细胞的体外测定中确定给定的AUG碱基三联体是否具有开始翻译的潜力:将特征在于去除了天然起始密码子并且在天然起始密码子的去除位置下游包含给定的AUG碱基三联体的RNA复制子引入宿主细胞。在一个实施方案中,宿主细胞是来自真核物种的细胞,所述真核物种是包含天然甲病毒5'复制识别序列的特定甲病毒的天然宿主。在一个优选实施方案中,宿主细胞是来自细胞系“BHK21[C13](ATCC® CCL10™)”的BHK21细胞,可从美国弗吉尼亚州马纳萨斯美国典型培养物保藏中心获得。优选地,在天然起始密码子的去除位置和给定的AUG碱基三联体之间不存在另外的AUG碱基三联体。如果在将去除了天然起始密码子并且包含给定的AUG碱基三联体为特征的RNA复制子转移到宿主细胞中之后在给定的AUG碱基三联体处开始翻译,则确定了给定的AUG碱基三联体具有开始翻译的潜力。可通过本领域已知的任何合适的方法来确定是否开始翻译。例如,复制子可在给定的AUG基础三联体的下游并且与给定的AUG碱基三联体同框地编码有利于检测翻译产物(如果有的话)的标签,例如myc-标签或HA-标签;可例如通过Western印迹确定是否存在具有编码的标签的表达产物。在该实施方案中,优选在给定的AUG碱基三联体和编码标签的核酸序列之间不存在其他AUG碱基三联体。可对多于一个给定的AUG碱基三联体分别进行基于细胞的体外测定:在每种情况下,优选在天然起始密码子的去除位置和给定的AUG碱基三联体之间不存在其他的AUG碱基三联体。这可通过去除天然起始密码子的去除位置和给定的AUG碱基三联体之间的所有AUG碱基三联体(如果有的话)来实现。因此,给定的AUG碱基三联体是天然起始密码子的去除位置下游的第一AUG碱基三联体。

[0273] 优选地,根据本发明的RNA复制子的5'复制识别序列的特征在于去除了所有潜在的起始密码子。因此,根据本发明,5'复制识别序列优选不包含可翻译成蛋白质的开放阅读框。

[0274] 在一个实施方案中,根据本发明的RNA复制子的5'复制识别序列的特征在于二级结构,其等同于病毒基因组RNA的5'复制识别序列的(预测的)二级结构。为此,RNA复制子可包含一个或更多个核苷酸变化,以补偿由去除至少一个起始密码子而引入的一个或更多个茎环内的核苷酸配对破坏。

[0275] 在一个实施方案中,根据本发明的RNA复制子的5'复制识别序列的特征在于与甲病毒基因组RNA的5'复制识别序列的二级结构等同的二级结构。在一个优选实施方案中,根据本发明的RNA复制子的5'复制识别序列的特征在于与甲病毒基因组RNA的5'复制识别序列的预测的二级结构等同的预测的二级结构。根据本发明,RNA分子的二级结构优选地由用

于RNA二级结构预测的网络服务器预测, <http://ma.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html>。

[0276] 通过比较以与天然甲病毒5'复制识别序列相比去除了至少一个起始密码子为特征的RNA复制子的5'复制识别序列的二级结构或预测的二级结构,可确定核苷酸配对破坏的存在或不存在。例如,与天然甲病毒5'复制识别序列相比,给定位置处至少一个碱基对可能不存在,例如茎环中的,特别是茎环的茎中的碱基对。

[0277] 在一个实施方案中,5'复制识别序列的一个或更多个茎环不被缺失或破坏。更优选地,茎环3和4不被缺失或破坏。优选地,5'复制识别序列的茎环都不被缺失或破坏。

[0278] 在一个实施方案中,至少一个起始密码子的去除不破坏5'复制识别序列的二级结构。在一个替代实施方案中,至少一个起始密码子的去除确实破坏了5'复制识别序列的二级结构。在该实施方案中,与天然5'复制识别序列相比,至少一个起始密码子的去除可以是在给定位置处的至少一个碱基对(例如茎环中的碱基对)不存在的原因。与天然5'复制识别序列相比,如果茎环中的碱基对不存在,则至少一个起始密码子的去除被确定为在茎环内引入了核苷酸配对破坏。茎环中的碱基对通常是茎环的茎中的碱基对。

[0279] 在一个实施方案中, RNA复制子包含一个或更多个核苷酸变化以补偿通过去除至少一个起始密码子而引入的一个或更多个茎环内的核苷酸配对破坏。

[0280] 如果与天然5'复制识别序列相比,至少一个起始密码子的去除在茎环内引入了核苷酸配对破坏,则可引入预期补偿核苷酸配对破坏的一个或更多个核苷酸变化,并且可由此获得的二级结构或预测的二级结构与天然5'复制识别序列进行比较。

[0281] 基于公知常识和本文中的公开内容,技术人员可预期某些核苷酸变化以补偿核苷酸配对破坏。例如,与天然5'复制识别序列相比,如果以至少一个起始密码子的去除为特征的RNA复制子的给定5'复制识别序列的二级结构或预测的二级结构的给定位置处的碱基对被破坏,则恢复该位置处的碱基对,优选不重新引入起始密码子的核苷酸变化预期将补偿核苷酸配对破坏。

[0282] 在一个实施方案中,复制子的5'复制识别序列不与可翻译核酸序列(即可翻译成肽或蛋白质,特别是nsP,特别是nsP1,或其任一个的片段)重叠或不包含该可翻译核酸序列。对于“可翻译”的核苷酸序列,其需要存在起始密码子;起始密码子编码肽或蛋白质的最N端的氨基酸残基。在一个实施方案中,复制子的5'复制识别序列不与编码nsP1的N端片段的可翻译核酸序列重叠或不包含该可翻译核酸序列。

[0283] 在一些情况下, RNA复制子包含至少一个亚基因组启动子。在一个优选实施方案中,复制子的亚基因组启动子不与可翻译核酸序列(即可翻译成肽或蛋白质,特别是nsP,特别是nsP4,或其任一个的片段)重叠或不包含该可翻译核酸序列。在一个实施方案中,复制子的亚基因组启动子不与编码nsP4的C端片段的可翻译核酸序列重叠或不包含该可翻译核酸序列。通过缺失nsP4的编码序列的一部分(通常是编码nsP4的N端部分的部分)和/或通过去除nsP4编码序列的未缺失的部分中的AUG碱基三联体,可产生具有亚基因组启动子的RNA复制子,所述亚基因组启动子不与例如可翻译成nsP4的C端片段的可翻译核酸序列重叠或不包含该可翻译核酸序列。如果去除nsP4或其一部分的编码序列中的AUG碱基三联体,则去除的AUG碱基三联体优选是潜在的起始密码子。或者,如果亚基因组启动子不与编码nsP4的核酸序列重叠,则可缺失编码nsP4的整个核酸序列。

[0284] 在一个实施方案中, RNA复制子不包含编码截短的非结构蛋白例如截短的甲病毒非结构蛋白的开放阅读框。在该实施方案的上下文中, 特别优选地, RNA复制子不包含编码nsP1的N端片段的开放阅读框, 并且任选地不包含编码nsP4的C端片段的开放阅读框。nsP1的N端片段是截短的甲病毒蛋白; nsP4的C端片段也是截短的甲病毒蛋白。

[0285] 在一些实施方案中, 根据本发明的复制子不包含甲病毒基因组5'端的茎环2(SL2)。根据Frolov等人(上文), 茎环2是存在于甲病毒基因组的5'端处在CSE 2上游的保守二级结构, 但对于复制是非必需的。

[0286] 根据本发明的RNA复制子优选是单链RNA分子。根据本发明的RNA复制子通常是(+)链RNA分子。在一个实施方案中, 本发明的RNA复制子是分离的核酸分子。根据本发明的RNA复制子包含至少一个经修饰核苷酸, 并且优选包含一个或更多个序列变化, 特别是通过本文中公开的方法所检测到的序列变化, 所述方法用于鉴定恢复或提高包含至少一个经修饰核苷酸的rRNA功能的序列变化。

[0287] 至少一个编码至少一种目的基因产物的开放阅读框

[0288] 在一个实施方案中, 根据本发明的RNA复制子包含至少一个编码目的基因产物(例如目的肽或目的蛋白质)的开放阅读框。优选地, 目的蛋白质由异源核酸序列编码。编码目的肽或目的蛋白质的基因同义地称为“目的基因”或“转基因”。在多个实施方案中, 目的肽或目的蛋白质由异源核酸序列编码。根据本发明, 术语“异源”是指核酸序列不是天然功能上或结构上与病毒核酸序列例如甲病毒核酸序列连接的事实。

[0289] 根据本发明的复制子可编码单一多肽或多个多肽。多个多肽可编码为单个多肽(融合多肽)或分开的多肽。在一些实施方案中, 根据本发明的复制子可包含多于一个开放阅读框, 每个开放阅读框可独立地选择为在亚基因组启动子的控制之下或不在其控制下。或者, 多蛋白或融合多肽包含由2A自切割肽(例如来自口蹄疫病毒2A蛋白)或蛋白酶切割位点或内含肽所分隔的单独的多肽。

[0290] 目的蛋白质可例如选自报道蛋白、药学活性肽或药学活性蛋白质、胞内干扰素(interferon, IFN)信号传导的抑制剂。根据本发明, 目的蛋白质优选地不包括来自自复制病毒的功能性非结构蛋白, 例如功能性甲病毒的非结构蛋白。

[0291] 报道蛋白

[0292] 在一个实施方案中, 开放阅读框编码报道蛋白, 例如细胞表面表达的蛋白例如CD90。在该实施方案中, 开放阅读框包含报道基因。可选择某些基因作为报道子, 因为这些基因赋予表达其的细胞或生物体的特征可容易地鉴定和测量, 或者因为这些基因是可选择的标志物。报道基因通常用作指示某个基因是否已被细胞或生物体群体吸收或者在细胞或生物体群体中表达。优选地, 报道基因的表达产物是视觉上可检测的。常见的视觉上可检测的报道蛋白通常具有荧光或发光蛋白。特定报道基因的实例包括编码水母绿色荧光蛋白(green fluorescent protein, GFP)、萤光素酶和红色荧光蛋白(red fluorescent protein, RFP)的基因, 所述水母绿色荧光蛋白导致表达其的细胞在蓝光下发光绿色, 所述萤光素酶催化与萤光素的反应以产生光。这些特定报道基因中任一者的变体都是可能的, 只要该变体具有视觉上可检测的特性即可。例如, eGFP是GFP的点突变体变体。报道蛋白实施方案特别适合于测试表达。

[0293] 药学活性基因产物, 例如肽或蛋白质或核酸

[0294] 根据本发明,在一个实施方案中,rRNA包含药学活性rRNA或由药学活性rRNA组成。“药学活性RNA”可以是编码药学活性肽或药学活性蛋白质的RNA。优选地,根据本发明的RNA复制子编码药学活性肽或药学活性蛋白质或者其他基因产物。优选地,开放阅读框编码药学活性肽或药学活性蛋白质。优选地,RNA复制子包含编码药学活性肽或药学活性蛋白质的开放阅读框,其任选地在亚基因组启动子的控制下。

[0295] 当“药学活性肽或药学活性蛋白质”以治疗有效量向对象施用,其对对象的病症或疾病状态具有积极或有利的的作用。优选地,药学活性肽或药学活性蛋白质具有治愈或姑息特性,并且可被施用以改善、减轻、缓解、逆转疾病或病症的一种或更多种症状,延迟疾病或病症的一种或更多种症状的发作或减轻其严重性。药学活性肽或药学活性蛋白质可具有预防特性,并可用于延迟疾病的发作或者减轻这样的疾病或病理状况的严重性。术语“药学活性肽或药学活性蛋白质”包括完整的蛋白质或多肽,并且还可以是指其药学活性片段。其还可包括肽或蛋白质的药学活性类似物。术语“药学活性肽或药学活性蛋白质”包括是抗原的肽和蛋白质,即肽或蛋白质在对象中引发免疫应答,其可以是治疗性或者部分或完全防止性的。

[0296] 在一个实施方案中,药学活性肽或药学活性蛋白质是或者包含免疫活性化合物或抗原或表位。

[0297] 根据本发明,术语“免疫活性化合物”涉及优选通过诱导和/或抑制免疫细胞的成熟、诱导和/或抑制细胞因子生物合成来改变免疫应答以及/或者通过刺激B细胞产生抗体来改变体液免疫的任何化合物。在一个实施方案中,免疫应答涉及刺激抗体应答(通常包括免疫球蛋白G(immunoglobulin G,IgG))。免疫活性化合物具有有效的免疫刺激活性,包括但不限于抗病毒和抗肿瘤活性,并且还可下调免疫应答的其他方面,例如将免疫应答从Th2免疫应答转移,这对于治疗广泛范围的Th2介导的疾病是有用的。

[0298] 根据本发明,术语“抗原”或“免疫原”涵盖将引发免疫应答的任何物质。特别地,“抗原”涉及与抗体或T淋巴细胞(T细胞)特异性反应的任何物质。根据本发明,术语“抗原”包含含有至少一个表位的任何分子。优选地,本发明背景下的抗原是任选在加工之后诱导免疫反应的分子,其优选对抗原具有特异性。根据本发明,可使用任何合适的抗原,其是免疫反应的候选物,其中免疫反应可以是体液免疫反应和细胞免疫反应二者。在本发明的实施方案的背景下,在MHC分子的背景下,抗原优选由细胞呈递,优选由抗原呈递细胞呈递,其导致针对抗原的免疫反应。抗原优选是对应于或来源于天然存在的抗原的产物。这样的天然存在的抗原可包括或可来源于变应原、病毒、细菌、真菌、寄生虫和其他感染原和病原体,或者抗原也可以是肿瘤抗原。根据本发明,抗原可对应于天然存在的产物,例如病毒蛋白质,或其一部分。在一些优选实施方案中,抗原是表面多肽,即天然展示在细胞、病原体、细菌、病毒、真菌、寄生虫、变应原或肿瘤表面上的多肽。抗原可引发针对细胞、病原体、细菌、病毒、真菌、寄生虫、变应原或肿瘤的免疫应答。

[0299] 术语“病原体”是指能够在生物体,优选脊椎动物生物体中引起疾病的致病性生物材料。病原体包括微生物,例如细菌、单细胞真核生物(原生动物)、真菌以及病毒。

[0300] 术语“表位”、“抗原肽”、“抗原表位”、“免疫原性肽”和“MHC结合肽”在本文中可互换使用,并且是指分子例如抗原中的抗原决定簇,即指由免疫系统识别例如由T细胞识别的免疫活性化合物的一部分或片段,特别是当在MHC分子的背景下呈递时。蛋白质的表位优选

包含所述蛋白质的连续或不连续部分,并且长度优选5至100个氨基酸,优选5至50、更优选8至30、最优选10至25个氨基酸,例如,表位的长度可优选9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个氨基酸。根据本发明,表位可与MHC分子例如细胞表面上的MHC分子结合,并且因此可以是“MHC结合肽”或“抗原肽”。术语“主要组织相容性复合体”和缩写“MHC”包括MHC I类和MHC II类分子,并且涉及存在于所有脊椎动物中的基因的复合体。MHC蛋白或分子对于免疫反应中淋巴细胞和抗原呈递细胞或患病细胞之间的信号传导是重要的,其中MHC蛋白或分子结合肽并将其呈递以通过T细胞受体识别。由MHC编码的蛋白质在细胞表面上表达,并将自身抗原(来自细胞本身的肽片段)和非自身抗原(例如,入侵微生物的片段)二者向T细胞展示。优选这样的免疫原性部分结合至MHC I类或II类分子。如本文中所使用的,如果使用本领域已知的任何测定可检出这样的结合,则称免疫原性部分“结合至”MHC I类或II类分子。术语“MHC结合肽”涉及结合至MHC I类和/或MHC II类分子的肽。在I类MHC/肽复合体的情况下,尽管更长或更短的肽可以是有效的,但结合肽通常是8至10个氨基酸长。在II类MHC/肽复合体的情况下,结合肽通常是10至25个氨基酸长,并且特别是13至18个氨基酸长,而更长和更短的肽可以是有效的。

[0301] 在一个实施方案中,根据本发明的目的蛋白质包含适合于靶生物体的疫苗接种的表位。本领域技术人员将知晓,免疫生物学和疫苗接种的原理之一是基于这样的事实,即通过用抗原免疫生物体来产生对疾病的免疫保护反应,所述抗原与待治疗的疾病具有免疫学相关性。根据本发明,抗原选自自身抗原和非自身抗原。非自身抗原优选是细菌抗原、病毒抗原、真菌抗原、变应原或寄生虫抗原。优选的是抗原包含能够在靶生物体中引发免疫应答的表位。例如,表位可引发针对细菌、病毒、真菌、寄生虫、变应原或肿瘤的免疫应答。

[0302] 在一些实施方案中,非自身抗原是细菌抗原。在一些实施方案中,抗原引发针对感染动物(包括鸟、鱼和哺乳动物,包括驯养动物)的细菌的免疫应答。优选地,引发针对其的免疫应答的细菌是致病性细菌。

[0303] 在一些实施方案中,非自身抗原是病毒抗原。病毒抗原可以是例如来自病毒表面蛋白的肽,例如如来自冠状病毒的衣壳多肽或刺突多肽。在一些实施方案中,抗原引发针对感染动物(包括鸟、鱼和哺乳动物,包括驯养动物)的病毒的免疫应答。优选地,引发针对其的免疫应答的病毒是致病性病毒。

[0304] 在一些实施方案中,非自身抗原是来自真菌的多肽或蛋白质。在一些实施方案中,抗原引发针对感染动物(包括鸟类,鱼类和哺乳动物,包括驯养动物)的真菌的免疫应答。优选地,引发针对其的免疫应答的真菌是致病性真菌。

[0305] 在一些实施方案中,非自身抗原是来自单细胞真核寄生虫的多肽或蛋白质。在一些实施方案中,抗原引发针对单细胞真核寄生虫,优选致病性单细胞真核寄生虫的免疫应答。致病性单细胞真核寄生虫可以是例如来自疟原虫(*Plasmodium*)属,例如恶性疟原虫(*P.falciparum*)、间日疟原虫(*P.vivax*)、三日疟原虫(*P.malariae*)或卵形疟原虫(*P.ovale*),来自利什曼原虫(*Leishmania*)属,或者来自锥虫(*Trypanosoma*)属,例如克氏锥虫(*T.cruzi*)或布氏锥虫(*T.brucei*)。

[0306] 在一些实施方案中,非自身抗原是变应原性多肽或变应原性蛋白质。变应原性蛋白质或变应原性多肽适合于变应原免疫治疗,也称为脱敏作用(hypo-sensitization)。

[0307] 在一些实施方案中,抗原是自身抗原,特别是肿瘤抗原。本领域技术人员已知肿瘤

抗原及其测定。

[0308] 在本发明的背景下,术语“肿瘤抗原”或“肿瘤相关抗原”涉及在正常条件下在有限数量的组织和/或器官中或者在特定发育阶段中特异性表达的蛋白质,例如,肿瘤抗原可在正常条件下在胃组织中特异性表达,优选在胃黏膜中、在生殖器官中(例如在睾丸中)、在滋养层组织中(例如在胎盘中)或在种系细胞中表达,并且在一种或更多种肿瘤或癌组织中表达或异常表达。在这种背景下,“有限数量”优选地意指不多于3,更优选不多于2。本发明背景下的肿瘤抗原包括例如分化抗原,优选细胞类型特异性分化抗原,即在正常条件下在特定分化阶段下在特定细胞类型中特异性表达的蛋白质;癌症/睾丸抗原,即在正常条件下在睾丸并且有时在胎盘中特异性表达的蛋白质;以及种系特异性抗原。在本发明的背景下,肿瘤抗原优选与癌细胞的细胞表面相关,并且优选在正常组织中不表达或仅很少表达。优选地,肿瘤抗原或肿瘤抗原的异常表达可鉴定癌细胞。在本发明的背景下,由对象(例如患有癌症疾病的患者)中的癌细胞表达的肿瘤抗原,优选是所述对象中的自身蛋白质。在一些优选实施方案中,本发明背景下的肿瘤抗原在正常条件下在非必需的组织或器官(即当被免疫系统损伤时不导致对象死亡的组织或器官)中,或者在免疫系统不能或仅几乎不能到达的身体器官或结构中特异性表达。优选地,肿瘤抗原的氨基酸序列在正常组织中表达的肿瘤抗原与在癌组织中表达的肿瘤抗原之间是相同的。

[0309] 可用于本发明中的肿瘤抗原的实例为p53、ART-4、BAGE、 β -联蛋白/m、Bcr-abLCAMEL、CAP-1、CASP-8、CDC27/m、CDK4/m、CEA、密蛋白家族的细胞表面蛋白(例如密蛋白-6、密蛋白-18.2和密蛋白-12)、c-MYC、CT、Cyp-B、DAM、ELF2M、ETV6-AML1、G250、GAGE、GnT-V、Gap100、HAGE、HER-2/neu、HPV-E7、HPV-E6、HAST-2、hTERT(或hTRT)、LAGE、LDLR/FUT、MAGE-A(优选MAGE-A1、MAGE-A2、MAGE-A3、MAGE-A4、MAGE-A5、MAGE-A6、MAGE-A7、MAGE-A8、MAGE-A9、MAGE-A10、MAGE-A11或MAGE-A12)、MAGE-B、MAGE-C、MART-1/Melan-A、MC1R、肌球蛋白/m、MUC1、MUM-1、-2、-3、NA88-A、NF1、NY-ESO-1、NY-BR-1、p190 minor BCR-abL、Pm1/RARa、PRAME、蛋白酶3、PSA、PSM、RAGE、RU1或RU2、SAGE、SART-1或SART-3、SCGB3A2、SCP1、SCP2、SCP3、SSX、存活蛋白、TEL/AML1、TPI/m、TRP-1、TRP-2、TRP-2/INT2、TPTE和WT。一些特别优选的肿瘤抗原包括密蛋白-18.2(CLAUDIN-18.2,CLDN18.2)和密蛋白-6(CLAUDIN-6,CLDN6)。

[0310] 在一些实施方案中,不要求药理学活性肽或药理学活性蛋白质是引发免疫应答的抗原。合适的药理学活性蛋白质或药理学活性肽可选自细胞因子和免疫系统蛋白质,例如免疫活性化合物(例如,白介素、集落刺激因子(colony stimulating factor,CSF)、粒细胞集落刺激因子(granulocyte colony stimulating factor,G-CSF)、粒细胞-巨噬细胞集落刺激因子(granulocyte-macrophage colony stimulating factor,GM-CSF)、促红细胞生成素、肿瘤坏死因子(tumor necrosis factor,TNF)、干扰素、整合素、地址素、选择素、归巢受体、T细胞受体、嵌合抗原受体(chimeric antigen receptor,CAR)、免疫球蛋白)、激素(胰岛素、甲状腺激素、儿茶酚胺、促性腺激素、促激素、催乳素、催产素、多巴胺、牛促生长素、瘦素等)、生长激素(例如,人生长激素)、生长因子(例如,表皮生长因子、神经生长因子、胰岛素样生长因子等)、生长因子受体、酶(组织纤溶酶原激活物、链激酶、胆固醇生物合成酶或降解酶、类固醇生成酶、激酶、磷酸二酯酶、甲基化酶、去甲基化酶、脱氢酶、纤维素酶、蛋白酶、脂肪酶、磷脂酶、芳香酶、细胞色素、腺苷酸或鸟苷酸环化酶、神经酰胺酶等)、受体(类固醇激素受体、肽受体)、结合蛋白(生长激素或生长因子结合蛋白等)、转录和翻译因子、肿瘤生

长抑制蛋白(例如,抑制血管生成的蛋白质)、结构蛋白(例如胶原、丝心蛋白、纤维蛋白原、弹性蛋白、微管蛋白、肌动蛋白和肌球蛋白)、血液蛋白(凝血酶、血清白蛋白、因子VII、因子VIII、胰岛素、因子IX、因子X、组织纤溶酶原激活物、蛋白C、von Willebrand因子、抗凝血酶III、葡萄糖苷脂酶、促红细胞生成素、粒细胞集落刺激因子(GCSF)或经修饰的因子VIII、抗凝剂等。在一个实施方案中,根据本发明的药学活性蛋白质是参与调节淋巴稳态(lymphoid homeostasis)的细胞因子,优选的是参与并优选地诱导或增强T细胞的发育、致敏、扩增、分化和/或存活的细胞因子。在一个实施方案中,细胞因子是白介素,例如IL-2、IL-7、IL-12、IL-15或IL-21。

[0311] 干扰素(IFN)信号传导的抑制剂

[0312] 由开放阅读框编码的另一合适的目的蛋白质是干扰素(IFN)信号传导的抑制剂。虽然已经报道了引入RNA以表达的细胞的生存力可能降低,特别是如果用RNA多次转染细胞的话,但是发现IFN抑制剂增强了其中将要表达RNA的细胞的生存力(WO 2014/071963 A1)。优选地,抑制剂是I型IFN信号传导的抑制剂。防止细胞外IFN与IFN受体的结合和抑制细胞中的细胞内IFN信号传导可使RNA在细胞中稳定表达。作为补充或替代,防止细胞外IFN与IFN受体的结合和抑制细胞内IFN信号传导可提高细胞的存活,特别是如果用RNA重复转染细胞的话。不希望受理论的束缚,设想细胞内IFN信号传导可导致翻译和/或RNA降解的抑制。这可通过抑制一种或更多种IFN诱导型抗病毒活性效应蛋白来解决。IFN诱导型抗病毒活性效应蛋白可选自RNA依赖性蛋白激酶(PKR)、2',5'-寡腺苷酸合成酶(OAS)和RNaseL。抑制细胞内IFN信号传导可包括抑制PKR依赖性途径和/或OAS依赖性途径。合适的目的蛋白质是能够抑制PKR依赖性途径和/或OAS依赖性途径的蛋白质。抑制PKR依赖性途径可包括抑制eIF2- α 磷酸化。抑制PKR可包括用至少一种PKR抑制剂处理细胞。PKR抑制剂可以是PKR的病毒抑制剂。PKR的优选的病毒抑制剂是痘苗病毒E3。如果肽或蛋白质(例如E3、K3)抑制细胞内IFN信号传导,则肽或蛋白质的细胞内表达是优选的。痘苗病毒E3是25kDa的dsRNA结合蛋白(由基因E3L编码),其结合并隔离dsRNA,以阻止PKR和OAS的激活。E3可直接与PKR结合并抑制其活性,从而导致eIF2- α 的磷酸化降低。IFN信号传导的其他合适的抑制剂是单纯疱疹病毒(Herpes simplex virus)ICP34.5、流感病毒(Influenza virus)NS1、托斯卡纳病毒(Toscana virus)NSs、家蚕核多角体病毒(Bombyx mori nucleopolyhedrovirus)PK2和HCV NS3A。

[0313] rRNA分子中至少一个开放阅读框的位置

[0314] rRNA复制子适合于任选在亚基因组启动子的控制下表达一种或更多种编码目的肽或目的蛋白质的基因。多个实施方案是可能的。RNA复制子上可存在一个或更多个开放阅读框,每个开放阅读框编码目的肽或目的蛋白质。RNA复制子的最上游开放阅读框称为“第一开放阅读框”。在一个实施方案中,编码目的蛋白质的第一开放阅读框位于5'复制识别序列的下游并且位于IRES的上游(并且开放阅读框编码来自自复制病毒的功能性非结构蛋白)。在一些实施方案中,“第一开放阅读框”是RNA复制子的唯一开放阅读框。任选地,在第一开放阅读框的下游可存在一个或更多个另外的开放阅读框。第一开放阅读框下游的一个或更多个另外的开放阅读框以其存在于第一开放阅读框下游的顺序(5'至3')可称为“第二开放阅读框”、“第三开放阅读框”,依此类推。在一个实施方案中,编码一个或更多个目的蛋白质的一个或更多个另外的开放阅读框位于编码来自自复制病毒的功能性非结构蛋白的

开放阅读框的下游,并且优选地由亚基因组启动子控制。优选地,每个开放阅读框包含起始密码子(碱基三联体),通常为AUG(在RNA分子中),对应于ATG(在各自的DNA分子中)。

[0315] 如果复制子包含3'复制识别序列,则优选地所有开放阅读框都位于3'复制识别序列的上游。

[0316] 在一些实施方案中,复制子的至少一个开放阅读框在亚基因组启动子,优选甲病毒亚基因组启动子的控制下。甲病毒亚基因组启动子非常有效,并且因此适合于高水平的异源基因表达。优选地,亚基因组启动子是甲病毒中亚基因组转录物的启动子。这表示亚基因组启动子是对于甲病毒天然的并且优选控制所述甲病毒中开放阅读框转录的启动子,所述开放阅读框编码一种或更多种结构蛋白。或者,亚基因组启动子是甲病毒的亚基因组启动子的变体;充当宿主细胞中亚基因组RNA转录的启动子的任何变体都是合适的。如果复制子包含亚基因组启动子,则优选地复制子包含保守序列元件3(CSE 3)或其变体。

[0317] 优选地,在亚基因组启动子控制下的至少一个开放阅读框位于亚基因组启动子的下游。优选地,亚基因组启动子控制包含开放阅读框的转录物的亚基因组RNA的产生。

[0318] 在一些实施方案中,第一开放阅读框在亚基因组启动子的控制下。在一个实施方案中,当第一开放阅读框处于亚基因组启动子的控制下时,由第一开放阅读框编码的基因既可从复制子中表达,也可从其亚基因组转录物中表达(后者在存在功能性甲病毒非结构蛋白的情况下)。分别处于亚基因组启动子控制下的一个或更多个另外的开放阅读框可存在于可在亚基因组启动子控制下的第一开放阅读框的下游。由一个或更多个另外的开放阅读框(例如由第二开放阅读框)编码的基因可从分别处于亚基因组启动子的控制下的一个或更多个亚基因组转录物中翻译。例如,RNA复制子可包含亚基因组启动子,其控制编码第二目的蛋白质的转录物的产生。

[0319] 在另一些实施方案中,第一开放阅读框不受亚基因组启动子的控制。在一个实施方案中,当第一开放阅读框不受亚基因组启动子控制时,由第一开放阅读框编码的基因可从复制子表达。分别处于亚基因组启动子的控制下的一个或更多个另外的开放阅读框可存在于第一开放阅读框的下游。可从亚基因组转录物表达由一个或更多个另外的开放阅读框编码的基因。

[0320] 在包含根据本发明的复制子的细胞中,复制子可通过功能性非结构蛋白扩增。另外,如果复制子包含在亚基因组启动子的控制下的一个或更多个开放阅读框,则预期由功能性非结构蛋白来制备一种或更多种亚基因组转录物。

[0321] 如果复制子包含多于一个编码目的蛋白质的开放阅读框,则优选每个开放阅读框编码不同的蛋白质。例如,由第二开放阅读框编码的蛋白质不同于由第一开放阅读框编码的蛋白质。

[0322] 根据本发明的可复制RNA分子的其他特征

[0323] 根据本发明的RNA分子可任选地以其他特征为特征,例如,5'帽、5'-UTR、3'-UTR、poly(A)序列和/或用于优化RNA分子翻译和/或稳定的密码子使用的适应性,如以下所详述的。

[0324] 帽

[0325] 在一些实施方案中,根据本发明的复制子包含5'帽。

[0326] 术语“5'帽”、“帽”、“5'帽结构”、“帽结构”同义使用,是指在存在于某些真核初级

转录物(例如前体信使RNA)的5'末端的二核苷酸。5'帽是这样的结构,其中(任选经修饰的)鸟苷通过5'至5'三磷酸键(或在某些帽类似物的情况下经修饰的三磷酸键)结合至mRNA分子的第一核苷酸。该术语可以是指常规帽或帽类似物。

[0327] “包含5'帽的RNA”或“提供有5'帽的RNA”或“经5'帽修饰的RNA”或“加帽的RNA”是指包含5'帽的RNA。例如,提供具有5'帽的RNA可通过在所述5'帽的存在下DNA模板的体外转录来实现,其中所述5'帽被共转录并入到产生的RNA链中,或者可例如通过体外转录产生RNA,并且可使用加帽酶(例如痘苗病毒的加帽酶)将5'帽在转录后连接至RNA。在加帽RNA中,(加帽)RNA分子第一碱基的3'位置通过磷酸二酯键与RNA分子的后续碱基(“第二碱基”)的5'位置连接。

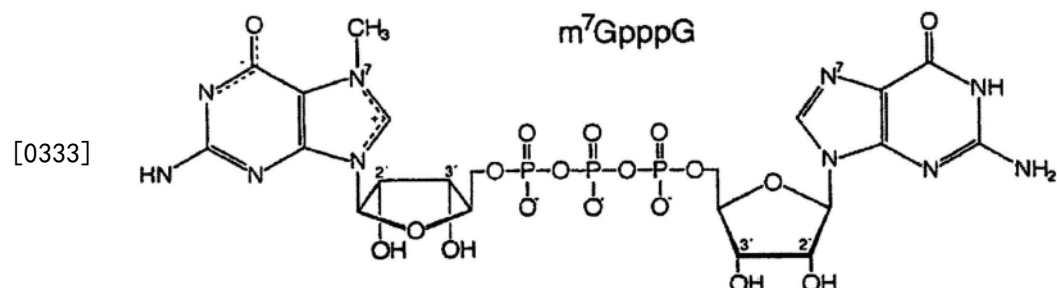
[0328] 在一个实施方案中,RNA复制子包含5'帽。在一个实施方案中,RNA复制子不包含5'帽。

[0329] 术语“常规5'帽”是指天然存在的5'帽,优选7-甲基鸟苷帽。在7-甲基鸟苷帽中,帽的鸟苷是经修饰的鸟苷,其中修饰由7-位的甲基化组成。

[0330] 在本发明的上下文中,术语“5'帽类似物”是指类似于常规5'帽但是被修饰以具有稳定RNA的能力(如果连接至其的话,优选在体内或细胞中)的分子结构。帽类似物不是常规5'帽。

[0331] 对于真核mRNA的情况,5'帽已被通常描述为参与mRNA的有效翻译:通常而言,在真核生物中,翻译仅在信使RNA(mRNA)分子的5'末端处开始,除非存在内部核糖体进入位点(IRES)。真核细胞能够在细胞核的转录过程中提供具有5'帽的RNA:新合成的mRNA通常被修饰具有5'帽结构,例如,当转录物达到20至30个核苷酸的长度时。首先,通过具有RNA5'-三磷酸酶和鸟苷酸转移酶活性的加帽酶将5'端核苷酸pppN(ppp代表三磷酸;N代表任何核苷)在细胞中转化为5'GpppN。GpppN随后可在细胞中由具有(鸟嘌呤-7)-甲基转移酶活性的第二种酶甲基化,以形成单甲基化的 m^7 GpppN帽。在一个实施方案中,本发明中使用的5'帽是天然5'帽。

[0332] 在本发明中,天然5'帽二核苷酸通常选自非甲基化的帽二核苷酸(G(5')ppp(5')N;也称为GpppN)和甲基化的帽二核苷酸($(m^7G(5'))ppp(5')N$;也称为 m^7 GpppN)。 m^7 GpppN(其中N为G)由以下式表示:



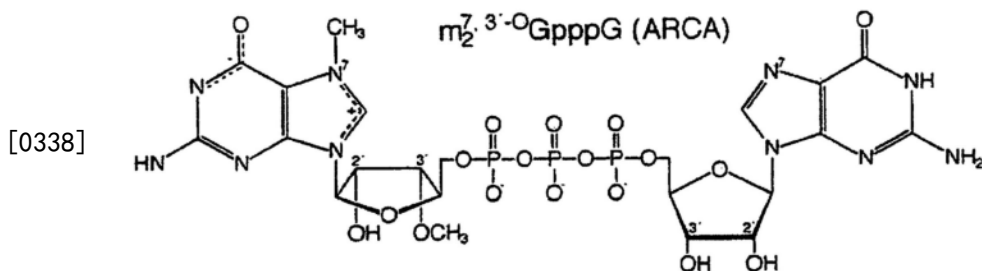
[0334] 本发明的加帽RNA可在体外制备,并且因此不依赖于宿主细胞中的加帽机制。在体外制备加帽RNA的最经常使用的方法是在存在所有四种核糖核苷三磷酸和帽二核苷酸(例如 $m^7G(5')ppp(5')G$ (也称为 m^7 GpppG))的情况下,用细菌或噬菌体任一者的RNA聚合酶转录DNA模板。RNA聚合酶通过 m^7 GpppG的鸟苷部分的3'-OH对下一个模板核苷三磷酸(pppN)的 α -磷酸的亲核攻击而开始转录,产生中间体 m^7 GpppGpN(其中N是RNA分子的第二碱基)。通过在体外转录过程中将帽与GTP的摩尔比设置为5至10,可抑制竞争性GTP起始产物pppGpN的形

成。

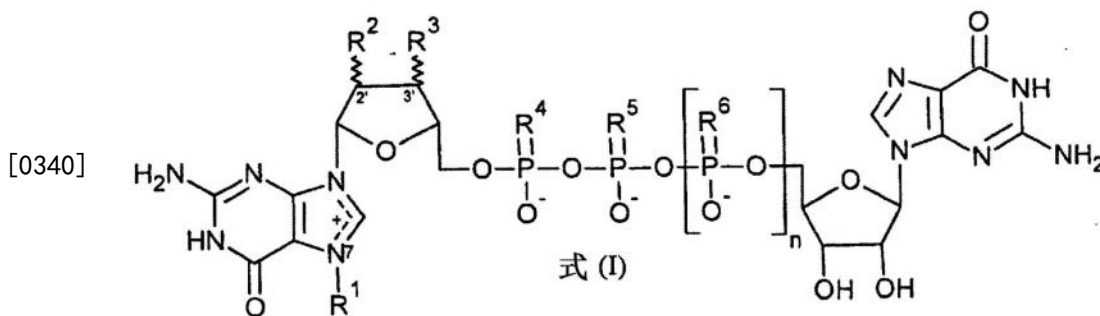
[0335] 在本发明的一些优选实施方案中,5'帽(如果存在)是5'帽类似物。如果RNA是通过体外转录获得的,例如是体外转录的RNA(IVT-RNA),则这些实施方案是特别合适的。最初已经描述了帽类似物通过体外转录来促进RNA转录物的大规模合成。

[0336] 对于信使RNA,迄今为止一般地已经描述了一些帽类似物(合成帽),并且其都可以在本发明的背景下使用。理想地,选择与较高的翻译效率和/或提高的体外降解抗性和/或提高的体内降解抗性相关的帽类似物。

[0337] 优选地,使用仅可以以一种方向并入RNA链的帽类似物。Pasquinelli et al., 1995, RNA J., 1:957-967)表明,在体外转录过程中,噬菌体RNA聚合酶使用7-甲基鸟苷单元来起始转录,其中约40%至50%的具有帽的转录物具有相反方向的帽二核苷酸(即初始反应产物为Gpppm⁷GpN)。与具有正确帽的RNA相比,具有反向帽的RNA在核酸序列翻译成蛋白质方面不起作用。因此,期望以正确的方向并入帽,即产生具有基本上对应于m⁷GpppGpN等的结构的RNA。已经表明,帽二核苷酸的反向整合被甲基化鸟苷单元的2'-或3'-OH基任一者的取代所抑制(Stepinski et al., 2001, RNA J. 7:1486-1495; Peng et al., 2002, Org. Lett. 24:161-164)。与存在常规5'帽m⁷GpppG的情况下体外转录的RNA相比,在存在这样的“抗反向帽类似物”的情况下合成的RNA更高效地翻译。为此,描述了一种帽类似物,其中甲基化鸟苷单元的3' OH基被OCH₃取代,例如Holtkamp et al., 2006, Blood, 108:4009-4017 (7-甲基(3'-O-甲基)GpppG; 抗反向帽类似物(anti-reverse cap analog, ARCA))。ARCA是根据本发明的合适的帽二核苷酸。



[0339] 在一个实施方案中,本发明的RNA基本上不易脱帽。这是重要的,因为通常而言,从引入到培养的哺乳动物细胞中的合成mRNA所产生的蛋白质的量受到mRNA天然降解的限制。一种用于mRNA降解的体内途径开始于mRNA帽的去除。这种去除由异二聚体焦磷酸酶催化,该酶含有调节亚基(Dcp1)和催化亚基(Dcp2)。催化亚基在三磷酸桥的 α 和 β 磷酸酯基团之间切割。在本发明中,可选择或存在对这种类型的切割不敏感或较不敏感的帽类似物。为此目的合适的帽类似物可选自根据式(I)的帽二核苷酸:



[0341] 其中R¹选自任选经取代的烷基、任选经取代的烯基、任选经取代的炔基、任选经取

代的环烷基、任选经取代的杂环基、任选经取代的芳基和任选经取代的杂芳基,

[0342] R^2 和 R^3 独立地选自H、卤素、OH和任选经取代的烷氧基,或者 R^2 和 R^3 一起形成O-X-O,其中X选自任选经取代的 CH_2 、 CH_2CH_2 、 $CH_2CH_2CH_2$ 、 $CH_2CH(CH_3)$,和

[0343] $C(CH_3)_2$,或者 R^2 与 R^3 连接至的环的4'位处的氢原子结合以形成-O- CH_2 -或- CH_2 -O-,

[0344] R^5 选自S、Se和 BH_3 ,

[0345] R^4 和 R^6 独立地选自O、S、Se和 BH_3 。

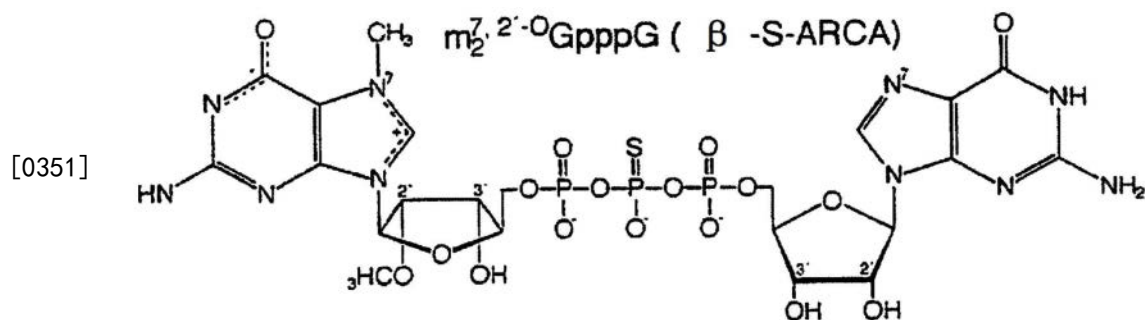
[0346] n为1、2或3。

[0347] R^1 、 R^2 、 R^3 、 R^4 、 R^5 、 R^6 的一些优选实施方案在WO 2011/015347 A1中公开,并且可在本发明中相应地选择。

[0348] 例如,在一个实施方案中,本发明的RNA包含硫代磷酸酯帽类似物。硫代磷酸酯-帽类似物是特定的帽类似物,其中三磷酸链中三个非桥连的O原子中的一者被S原子替换,即式(I)中的 R^4 、 R^5 或 R^6 中的一者是S。硫代磷酸酯-帽类似物已经由Kowalska et al., 2008, RNA, 14: 1119-1131描述为不期望的脱帽过程的解决方案,并且因此提高了RNA在体内的稳定性。特别地,在5'帽的 β -磷酸酯基团上用硫原子替换氧原子导致针对Dcp2的稳定。在本发明优选的该实施方案中,式(I)中的 R^5 为S;并且 R^4 和 R^6 为O。

[0349] 在另一个实施方案中,本发明的RNA包含硫代磷酸酯-帽类似物,其中RNA 5'帽的硫代磷酸酯修饰与“抗反向帽类似物”(ARCA)修饰组合。相应的ARCA-硫代磷酸酯-帽类似物在WO 2008/157688 A2中描述,并且其都可用于本发明的RNA中。在该实施方案中,式(I)中的 R^2 或 R^3 中的至少一个不是OH,优选地 R^2 和 R^3 中的一个为甲氧基(OCH_3),并且 R^2 和 R^3 中的另一个优选为OH。在一个优选的实施方案中, β -磷酸酯基团上氧原子被硫原子替换(使得式(I)中的 R^5 为S;并且 R^4 和 R^6 为O)。认为ARCA的硫代磷酸酯修饰确保了 α 、 β 和 γ 硫代磷酸酯基团在翻译和脱帽系统二者中都精确定位在帽结合蛋白的活性位点内。这些类似物中的至少一些基本对焦磷酸酶Dcp1/Dcp2具有抗性。硫代磷酸酯修饰的ARCA被描述为具有比缺少硫代磷酸酯基团的相应ARCA高得多的对eIF4E的亲和力。

[0350] 在本发明中特别优选的相应帽类似物(即 $m_2^{7,2'}\text{-}^0\text{GpppG}$)称为 β -S-ARCA(WO 2008/157688 A2;Kuhn et al., 2010, Gene Ther. 17:961-971)。因此,在本发明的一个实施方案中,本发明的RNA用 β -S-ARCA修饰。 β -S-ARCA由以下结构表示:



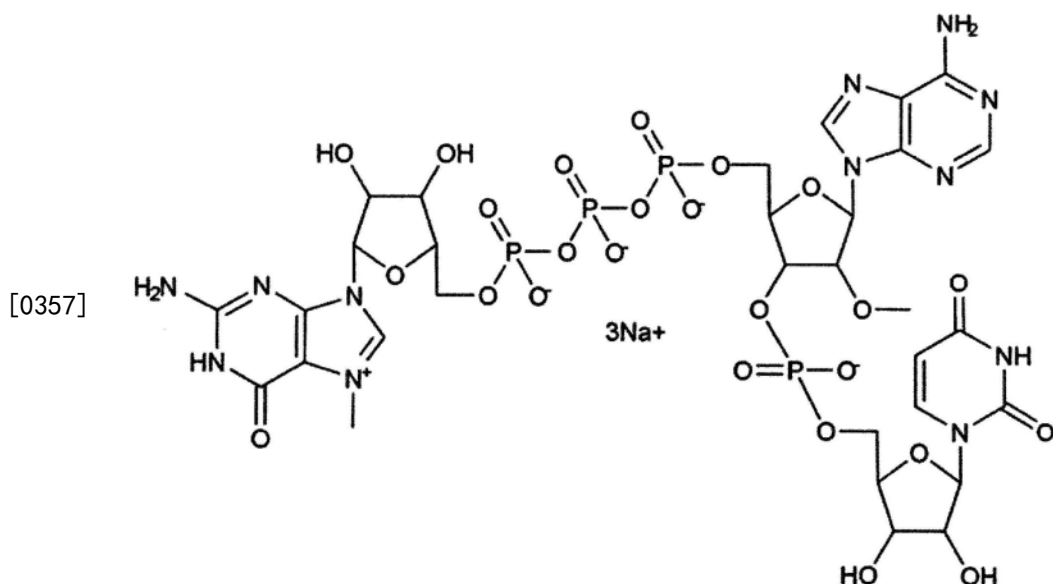
[0352] 通常而言,在桥接的磷酸酯处将氧原子替换为硫原子得到硫代磷酸酯非对映体,其基于在HPLC上的其洗脱模式称为D1和D2。简言之,“ β -S-ARCA的D1非对映体”或“ β -S-ARCA (D1)”是与 β -S-ARCA的D2非对映体(β -S-ARCA (D2))相比首先在HPLC柱上洗脱的 β -S-ARCA的非对映体,并且因此表现出较短的保留时间。WO 2011/015347 A1中描述了通过HPLC确定立体化学构型。

[0353] 在本发明的第一个特别优选的实施方案中,本发明的RNA用 β -S-ARCA (D2) 非对映体修饰。 β -S-ARCA的两种非对映体针对核酸酶的敏感性不同。已显示具有 β -S-ARCA的D2非对映体的RNA几乎完全抵抗Dcp2切割(与存在未经修饰的ARCA 5' 帽的情况下合成的RNA相比,只有6%的切割),而具有 β -S-ARCA (D1) 5' 帽的RNA对Dcp2切割表现出中等敏感性(71%切割)。进一步显示,针对Dcp2切割的稳定的提高与哺乳动物细胞中蛋白质表达的提高相关。特别地,已经显示,与具有 β -S-ARCA (D1) 帽的RNA相比,具有 β -S-ARCA (D2) 帽的RNA在哺乳动物细胞中更高效地翻译。因此,在本发明的一个实施方案中,本发明的RNA用根据式(I)的帽类似物修饰,其特征在于式(I)中包含取代基 R^5 的P原子处的立体化学构型,其对应于 β -S-ARCA的D2非对映体的 P_β 原子处的立体化学构型。在该实施方式中,式(I)中的 R^5 为S;并且 R^4 和 R^6 为O。另外,式(I)中的 R^2 或 R^3 中的至少一个优选不是OH,优选地 R^2 和 R^3 中的一个为甲氧基(OCH_3),并且 R^2 和 R^3 中的另一个优选为OH。

[0354] 在第二个特别优选的实施方案中,本发明的RNA用 β -S-ARCA (D1) 非对映体修饰。该实施方案特别适合于将加帽RNA转移到未成熟的抗原呈递细胞中,例如用于疫苗接种目的。已经表明,在将相应加帽的RNA转移到未成熟的抗原呈递细胞中时, β -S-ARCA (D1) 非对映体特别适合于提高RNA的稳定性,提高RNA的翻译效率,延长RNA的翻译,提高RNA的总蛋白质表达,和/或提高针对由所述RNA编码的抗原或抗原肽的免疫应答(Kuhn et al., 2010, Gene Ther. 17:961-971)。因此,在本发明的替代实施方案中,本发明的RNA用根据式(I)的帽类似物修饰,其特征在于式(I)中包含取代基 R^5 的P原子处的立体化学构型,其对应于 β -S-ARCA的D1非对映体的 P_β 原子处的立体化学构型。相应的帽类似物及其实施方案在WO 2011/015347 A1和Kuhn et al., 2010, Gene Ther. 17:961-971中描述。在WO 2011/015347 A1中描述任何帽类似物都可用于本发明,其中包含取代基 R^5 的P原子处的立体化学构型对应于 β -S-ARCA的D1非对映体的 P_β 原子处的立体化学构型。优选地,式(I)中的 R^5 为S;并且 R^4 和 R^6 为O。另外,式(I)中的 R^2 或 R^3 中的至少一个优选不是OH,优选地 R^2 和 R^3 中的一个为甲氧基(OCH_3),并且 R^2 和 R^3 中的另一个优选为OH。

[0355] 在一个实施方案中,本发明的RNA用根据式(I)的5' 帽结构修饰,其中任意一个磷酸酯基团被硼代磷酸酯基团或硒代磷酸酯基团取代。这样的帽在体外和体内二者均具有提高的稳定性。任选地,相应化合物具有2' -O-或3' -O-烷基(其中烷基优选为甲基);相应的帽类似物称为 BH_3 -ARCA或Se-ARCA。如WO 2009/149253 A2中所述,特别适合于mRNA加帽的化合物包括 β - BH_3 -ARCA和 β -Se-ARCA。对于这些化合物,优选地是在式(I)中包含取代基 R^5 的P原子处的立体化学构型对应于 β -S-ARCA的D1非对映体的 P_β 原子处的立体化学构型。

[0356] 在一个实施方案中,5' 帽可具有以下结构:



[0358] 在使用这种类型的帽的实施方案中,对应于甲病毒第二个核苷酸的U被排除在修饰之外,例如被排除在N1-甲基-假尿苷的修饰之外。

[0359] UTR

[0360] 术语“非翻译区”或“UTR”涉及DNA分子中转录但不翻译成氨基酸序列的区域,或者涉及RNA分子(例如mRNA分子)中的相应区域。非翻译区(UTR)可位于开放阅读框的5'(上游)(5'-UTR)和/或开放阅读框的3'(下游)(3'-UTR)。

[0361] 3'-UTR如果存在,则其位于基因的3'末端,在蛋白质编码区的终止密码子的下游,但是术语“3'-UTR”优选不包括poly(A)尾。因此,3'-UTR位于poly(A)尾(如果存在)的上游,例如紧邻poly(A)尾。

[0362] 5'-UTR如果存在,则其位于基因的5'末端,在蛋白质编码区的起始密码子的上游。5'-UTR位于5'帽(如果存在)的下游,例如紧邻5'帽。

[0363] 根据本发明,可将5'和/或3'非翻译区以这样的方式与开放阅读框功能性连接从而使这些区域与开放阅读框缔合:提高包含所述开放阅读框的RNA的稳定性和/或翻译效率。

[0364] 在一些实施方案中,根据本发明的RNA复制子包含5'-UTR和/或3'-UTR。

[0365] UTR与RNA的稳定性和翻译效率相关。除了涉及如本文中所述的5'帽和/或3' poly(A)尾的结构修饰外,通过选择特定的5'和/或3'非翻译区(untranslated region,UTR)可改善这二者。通常认为UTR中的序列元件会影响翻译效率(主要是5'-UTR)和RNA稳定性(主要是3'-UTR)。优选存在具有活性的5'-UTR,以提高RNA复制子的翻译效率和/或稳定性。独立地或另外地,优选存在具有活性的3'-UTR,以提高RNA复制子的翻译效率和/或稳定性。

[0366] 关于第一核酸序列(例如UTR)的术语“为了提高翻译效率而具有活性”和/或“为了提高稳定性而具有活性”意指第一核酸序列能够在与第二核酸序列的共同转录物中改变所述第二核酸序列的翻译效率和/或稳定性,这样的方式为与不存在所述第一核酸序列的情况下所述第二核酸序列的翻译效率和/或稳定性相比,所述翻译效率和/或稳定性提高。

[0367] 在一个实施方案中,根据本发明的RNA复制子包含对于功能性甲病毒非结构蛋白来源的甲病毒是异源或非天然的5'-UTR和/或3'-UTR。这允许根据所期望的翻译效率和RNA

稳定性来设计非翻译区。因此,异源或非天然UTR允许高度的灵活性,并且与天然甲病毒UTR相比,这种灵活性是有利的。

[0368] 优选地,根据本发明的RNA复制子包含不是病毒来源,特别不是甲病毒来源的5' -UTR和/或3' -UTR。在一个实施方案中,RNA复制子包含来源于真核5' -UTR的5' -UTR和/或来源于真核3' -UTR的3' -UTR。

[0369] 根据本发明的5' -UTR可包含任选地通过接头分开的多于一个核酸序列的任意组合。根据本发明的3' -UTR可包含任选地通过接头分开的多于一个核酸序列的任意组合。

[0370] 根据本发明的术语“接头”涉及添加在两个核酸序列之间以连接所述两个核酸序列的核酸序列。关于接头序列没有特别限制。

[0371] 3' -UTR通常具有200至2000个核苷酸的长度,例如500至1500个核苷酸。免疫球蛋白mRNA的3' 非翻译区相对较短(少于约300个核苷酸),而其他基因的3' 非翻译区相对较长。例如,tPA的3' -非翻译区的长度为约800个核苷酸,因子VIII的3' -非翻译区的长度为约1800个核苷酸,并且促红细胞生成素的3' -非翻译区的长度为约560个核苷酸。哺乳动物mRNA的3' 非翻译区通常具有被称为AAUAAA六核苷酸序列的同源区。该序列可以是poly (A)附着信号,并且通常位于poly (A)附着位点上游10至30个碱基处。3' -非翻译区可包含一个或更多个反向重复序列,其可折叠形成茎环结构,该结构充当外切核糖核酸酶的屏障或与已知提高RNA稳定性的蛋白质(例如RNA结合蛋白质)相互作用。

[0372] 人 β -珠蛋白3' -UTR,特别是人 β -珠蛋白3' -UTR的两个连续的相同拷贝有助于高的转录物稳定性和翻译效率(Holtkamp et al., 2006, Blood 108: 4009-4017)。因此,在一个实施方案中,根据本发明的RNA复制子包含人 β -珠蛋白3' -UTR的两个连续的相同拷贝。因此,其在5' $\rightarrow$ 3' 方向包含: (a) 任选的5' -UTR; (b) 开放阅读框; (c) 3' -UTR; 所述3' -UTR包含人 β -珠蛋白3' -UTR、其片段、或者人 β -珠蛋白3' -UTR的变体或其片段的两个连续的相同拷贝。

[0373] 在一个实施方案中,根据本发明的RNA复制子包含3' -UTR,其具有活性以提高翻译效率和/或稳定性,但该3' -UTR不是人 β -珠蛋白3' -UTR、其片段、或者人 β -珠蛋白3' -UTR的变体或其片段。

[0374] 在一个实施方案中,根据本发明的RNA复制子包含5' -UTR,其具有活性以提高翻译效率和/或稳定性。

[0375] Poly (A) 序列

[0376] 在一些实施方案中,根据本发明的复制子包含3' -poly (A) 序列。如果复制子包含保守序列元件4 (CSE 4),则复制子的3' -poly (A) 序列优选存在于CSE 4的下游,最优选紧邻CSE 4。

[0377] 根据本发明,在一个实施方案中,poly (A) 序列包含以下或基本上由以下组成或由以下组成:至少20个、优选至少26个、优选至少40个、优选至少80个、优选至少100个,并且优选至多500个、优选至多400个、优选至多300个、优选至多200个、并且特别是至多150个A核苷酸,并且特别是约120个A核苷酸。在本文中,“基本上由...组成”意指poly (A) 序列中的大多数核苷酸,通常按“poly (A) 序列”中核苷酸的数目计至少50%、并且优选至少75%是A核苷酸(腺苷酸),但允许其余核苷酸是A核苷酸以外的核苷酸,例如U核苷酸(尿苷酸)、G核苷酸(鸟苷酸)、C核苷酸(胞苷酸)。在本文中,“由...组成”意指poly (A) 序列中的所有核苷酸,即按poly (A) 序列中核苷酸的数目计100%是A核苷酸。术语“A核苷酸”或“A”是指腺苷酸。

[0378] 实际上,已表明约120个A核苷酸的3' poly(A)序列对转染的真核细胞中RNA的水平以及从存在于3' poly(A)序列的上游(5')的开放阅读框翻译的蛋白质的水平具有有益的影响(Holtkamp et al.,2006,Blood,vol.108,pp.4009-4017)。

[0379] 在甲病毒中,认为至少11个连续的腺苷酸残基或至少25个连续的腺苷酸残基的3' poly(A)序列对于有效合成负链很重要。特别是在甲病毒中,至少25个连续的腺苷酸残基的3' poly(A)序列被理解为与保守序列元件4(CSE 4)共同起作用以促进(-)链的合成(Hardy& Rice,2005,J.Virol.79:4630-4639)。

[0380] 本发明提供了RNA转录过程中(即体外转录的RNA的制备过程中)待连接的3' poly(A)序列,所述过程基于在与编码链互补的链中包含重复dT核苷酸(脱氧胸苷酸)的DNA模板。编码poly(A)序列的DNA序列(编码链)称为poly(A)盒。

[0381] 在本发明的一个优选实施方案中,存在于DNA的编码链中的3' poly(A)盒基本上由dA核苷酸组成,但是被具有四种核苷酸(dA、dC、dG、dT)的相等分布的随机序列中断。这样的随机序列的长度可以是5至50、优选10至30、更优选10至20个核苷酸。这样的盒在WO 2016/005004A1中公开。WO 2016/005004 A1中公开的任何poly(A)盒都可用于本发明。基本上由dA核苷酸组成但是被随机序列中断的poly(A)盒在DNA水平上显示质粒DNA在大肠杆菌(E.coli)中的恒定繁殖,并且在RNA水平上仍与关于支持RNA稳定性和翻译效率的有益特性相关,所述随机序列具有四种核苷酸(dA、dC、dG、dT)的相等分布并且长度为例如5至50个核苷酸。

[0382] 因此,在本发明的一个优选实施方案中,本文中所述的RNA分子中包含的3' poly(A)序列基本上由A核苷酸组成,但是被具有四种核苷酸(A、C、G、U)的相等分布的随机序列中断。这样的随机序列的长度可以是5至50、优选10至30、更优选10至20个核苷酸。

[0383] 密码子使用

[0384] 通常而言,遗传密码的简并性将允许RNA序列中存在的某些密码子(编码氨基酸的碱基三联体)由其他密码子(碱基三联体)替换,同时保持相同的编码能力(使得替换密码子与被替换密码子编码相同的氨基酸)。在本发明的一些实施方案中,由RNA(rRNA)分子包含的开放阅读框的至少一个密码子与该开放阅读框所起源的物种的相应开放阅读框中的相应密码子不同。在该实施方案中,开放阅读框的编码序列被称为是“适应的”或“修饰的”。可适应由复制子包含的开放阅读框的编码序列。

[0385] 例如,当适应开放阅读框的编码序列时,可选择通常使用的密码子:WO 2009/024567A1描述了核酸分子的编码序列的适应,涉及将稀有密码子替换为更通常使用的密码子。由于密码子使用的频率取决于宿主细胞或宿主生物体,因此这种适应类型适合使核酸序列适合在特定宿主细胞或宿主生物中表达。一般而言,尽管并不总是需要适应开放阅读框的所有密码子,但更通常使用的密码子通常在宿主细胞或宿主生物中更高效地翻译。

[0386] 例如,当适应开放阅读框的编码序列时,可通过为每个氨基酸选择具有最高富含GC含量的密码子来改变G(鸟苷酸)残基和C(胞苷酸)残基的含量。已经报道具有富含GC的开放阅读框的RNA分子具有降低免疫活化以及改善RNA翻译和半衰期的潜力(Thess et al.,2015,Mol.Ther.23:1457-1465)。

[0387] 特别地,非结构蛋白的编码序列可根据期望进行适应。由于编码非结构蛋白的开放阅读框不与复制子的5'复制识别序列重叠,这种自由是可能的。

[0388] 本发明实施方案的安全特性

[0389] 单独或以任何合适的组合的以下特性在本发明中是优选的:

[0390] 本发明的复制子不是形成颗粒的。这意指,在通过本发明的复制子接种宿主细胞之后,宿主细胞不产生病毒颗粒,例如下一代病毒颗粒。在一个实施方案中,根据本发明的RNA复制子完全不含编码任何病毒结构蛋白例如甲病毒结构蛋白(例如核心核衣壳蛋白C、包膜蛋白P62和/或包膜蛋白E1)的遗传信息。优选地,根据本发明的复制子不包含病毒包装信号,例如甲病毒包装信号。可去除例如包含在SFV的nsP2的编码区中的甲病毒包装信号(White et al.,1998,J.Virol.72:4320-4326),例如通过缺失或突变。去除甲病毒包装信号的合适方法包括nsP2编码区的密码子使用的适应。遗传密码的简并性可允许删除包装信号的功能,而不影响编码的nsP2的氨基酸序列。

[0391] DNA

[0392] 本发明还提供了DNA,其包含编码根据本发明的RNA复制子的核酸序列。

[0393] 优选地,DNA是双链的。

[0394] 在一个优选的实施方案中,DNA是质粒。如本文中使用的术语“质粒”通常涉及染色体外遗传物质的构建体,通常是环状DNA双链体,其可独立于染色体DNA复制。

[0395] 本发明的DNA可包含可被DNA依赖性RNA聚合酶识别的启动子。这允许编码的RNA(例如本发明的RNA)在体内或体外转录。IVT载体可以以标准化方式用作体外转录的模板。根据本发明优选的启动子的实例是SP6、T3或T7聚合酶的启动子。

[0396] 在一个实施方案中,本发明的DNA是分离的核酸分子。

[0397] 制备RNA的方法

[0398] 根据本发明的RNA分子可通过体外转录获得。体外转录的RNA(IVT-RNA)在本发明中特别令人感兴趣。IVT-RNA可通过从核酸分子(特别是DNA分子)转录获得。本发明的DNA分子适合于这样的目的,特别是如果包含可由DNA依赖性RNA聚合酶识别的启动子。

[0399] 根据本发明的rRNA可在体外合成。这允许在体外转录反应中添加帽类似物。通常而言,poly(A)尾由DNA模板上的poly-(dT)序列编码。或者,可在转录之后通过酶促实现加帽和poly(A)尾添加。

[0400] 体外转录方法是技术人员已知的。例如,如WO 2011/015347 A1中所述,多种体外转录试剂盒是可商购的。

[0401] 试剂盒

[0402] 本发明还提供了试剂盒,其包含根据本发明的RNA复制子。

[0403] 在一个实施方案中,试剂盒的成分作为单独的实体存在。例如,试剂盒的一种成分可存在于一个实体中,并且试剂盒的另一种成分可存在于单独的实体中。例如,打开或封闭的容器是合适的实体。封闭的容器是优选的。所使用的容器应优选无RNA酶或基本上无RNA酶。

[0404] 在一个实施方案中,本发明的试剂盒包含用于细胞接种和/或施用于人或动物对象的RNA。

[0405] 根据本发明的试剂盒任选地包含标签或其他形式的信息元素,例如电子数据载体。标签或信息元素优选地包含说明书,例如印刷的书面说明书或电子形式的说明书,其任选地可印刷。说明书可涉及试剂盒的至少一种合适的可能用途。

[0406] 药物组合物

[0407] 本文中所述的RNA复制子可以以药物组合物的形式存在。根据本发明的药物组合物可包含至少一种根据本发明的核酸分子。根据本发明的药物组合物包含可药用稀释剂和/或可药用赋形剂和/或可药用载体和/或可药用载剂。可药用载体、载剂、赋形剂或稀释剂的选择没有特别限制。可使用本领域已知的任何合适的可药用载体、载剂、赋形剂或稀释剂。

[0408] 在本发明的一个实施方案中,药物组合物还可包含溶剂,例如水性溶剂或任何使得可保留rRNA完整性的溶剂。在一个优选实施方案中,药物组合物是包含RNA的水溶液。水溶液可任选地包含溶质,例如盐。

[0409] 在本发明的一个实施方案中,药物组合物呈冻干组合物的形式。通过将相应的水性组合物进行冻干可获得冻干组合物。

[0410] 在一个实施方案中,药物组合物包含至少一种阳离子实体。通常而言,阳离子脂质、阳离子聚合物和其他带有正电荷的物质可与带负电荷的核酸形成复合物。通过与阳离子化合物,优选聚阳离子化合物,例如阳离子或聚阳离子肽或蛋白质复合,可使根据本发明的RNA稳定。在一个实施方案中,根据本发明的药物组合物包含选自鱼精蛋白、聚乙烯亚胺、聚-L-赖氨酸、聚-L-精氨酸、组蛋白或阳离子脂质的至少一种阳离子分子。

[0411] 根据本发明,阳离子脂质是阳离子两亲性分子,例如包含至少一个亲水性和亲脂性部分的分子。阳离子脂质可以是单阳离子或聚阳离子的。阳离子脂质通常具有亲脂性部分,例如固醇、酰基或二酰基链,并且具有总的净正电荷。脂质的头基通常带有正电荷。阳离子脂质优选具有1至10价的正电荷,更优选具有1至3价的正电荷,并且更优选具有1价的正电荷。阳离子脂质的实例包括但不限于1,2-二-0-十八碳烯基-3-三甲基铵丙烷(1,2-di-0-octadecenyl-3-trimethylammonium propane, DOTMA);二甲基双十八烷基铵(dimethyldioctadecylammonium, DDAB);1,2-二油酰基-3-三甲基铵丙烷(1,2-dioleoyl-3-trimethylammonium-propane, DOTAP);1,2-二油酰基-3-二甲基铵丙烷(1,2-dioleoyl-3-dimethylammonium-propane, DODAP);1,2-二酰氧基-3-二甲基铵丙烷;1,2-二烷氧基-3-二甲基铵丙烷;双十八烷基二甲基氯化铵(dioctadecyldimethyl ammonium chloride, DODAC)、1,2-二肉豆蔻酰氧基丙基-1,3-二甲基羟乙基铵(1,2-dimyristoyloxypropyl-1,3-dimethylhydroxyethyl ammonium, DMRIE)和2,3-二油酰氧基-N-[2(精胺羧酰胺)乙基]-N,N-二甲基-1-丙烷三氟乙酸盐(2,3-dioleoyloxy-N-[2(spermine carboxamide)ethyl]-N,N-dimethyl-1-propanamium trifluoroacetate, DOSPA)。阳离子脂质还包括具有叔胺基的脂质,包括1,2-二亚油基氧基-N,N-二甲基-3-氨基丙烷(1,2-dilinoleyloxy-N,N-dimethyl-3-aminopropane, DLinDMA)。阳离子脂质适合于如本文中所述的脂质制剂中配制RNA,例如脂质体、乳剂和脂质复合物。通常而言,正电荷由至少一种阳离子脂质贡献,并且负电荷由RNA贡献。在一个实施方案中,药物组合物除阳离子脂质外包含至少一种辅助脂质。辅助脂质可以是中性或阴离子脂质。辅助脂质可以是天然脂质,例如磷脂,或天然脂质的类似物,或完全合成的脂质,或与天然脂质没有相似性的类脂质分子。在药物组合物包含阳离子脂质和辅助脂质二者的情况下,可考虑制剂的稳定性等来适当确定阳离子脂质与中性脂质的摩尔比。

[0412] 在一个实施方案中,根据本发明的药物组合物包含鱼精蛋白。根据本发明,鱼精蛋

白可用作阳离子载体剂。术语“鱼精蛋白”是指多种相对低分子量的强碱性蛋白质中的任一种,其富含精氨酸,并且被发现特别是与DNA相关联,以代替动物(例如鱼)的精细胞中的体细胞组蛋白。特别地,术语“鱼精蛋白”是指存在于鱼精液中的蛋白质,其是强碱性的,可溶于水,不被热凝结并且包含多个精氨酸单体。根据本发明,如本文中使用的术语“鱼精蛋白”意指包括从天然或生物来源中获得或来源的任何鱼精蛋白氨基酸序列,所述氨基酸序列包括其片段以及所述氨基酸序列或其片段的多聚体形式。此外,该术语涵盖(合成的)多肽,这些多肽是人工的并且是为特定目的而专门设计的,并且不能从天然或生物来源中分离。

[0413] 在一些实施方案中,本发明的组合物可包含一种或更多种佐剂。可将佐剂添加至疫苗以刺激免疫系统的应答;佐剂本身通常不提供免疫力。示例性佐剂包括但不限于以下:无机化合物(例如明矾、氢氧化铝、磷酸铝、磷酸钙氢氧化物);矿物油(例如石蜡油)、细胞因子(例如IL-1、IL-2、IL-12);免疫刺激性多核苷酸(例如RNA或DNA;例如包含CpG的寡核苷酸);皂苷(例如来自皂树(*Quillaja*)、大豆、美远志(*Polygala senega*)的植物皂苷);油乳剂或脂质体;聚氧乙烯醚和聚氧乙烯酯制剂;聚磷腈(polyphosphazene,PCPP);胞壁酰肽;咪唑喹诺酮化合物;氨硫脲化合物;Flt3配体(WO 2010/066418 A1);或本领域技术人员已知的任何其他佐剂。根据本发明的用于施用RNA的一种优选佐剂是Flt3配体(WO 2010/066418 A1)。当Flt3配体与编码抗原的RNA一起施用,可观察到抗原特异性CD8⁺ T细胞的强的增加。

[0414] 可缓冲根据本发明的药物组合物(例如,用乙酸盐缓冲液、柠檬酸盐缓冲液、琥珀酸盐缓冲液、Tris缓冲液、磷酸盐缓冲液)。

[0415] 含RNA的颗粒

[0416] 在一些实施方案中,由于未保护的RNA的不稳定性,所以以复合或包封形式提供本发明的RNA分子是有利的。本发明提供了相应的药物组合物。特别地,在一些实施方案中,本发明的药物组合物包含含有核酸的颗粒,优选含有RNA的颗粒。相应的药物组合物称为颗粒制剂。在根据本发明的颗粒制剂中,颗粒包含根据本发明的核酸和适合于递送核酸的可药用载体或可药用载剂。含有核酸的颗粒可以是例如呈蛋白质颗粒的形式或呈含有脂质的颗粒的形式。合适的蛋白质或脂质称为颗粒形成剂。先前已经描述了蛋白质颗粒和含有脂质的颗粒适于以颗粒形式递送甲病毒RNA(例如Strauss&Strauss,1994,Microbiol.Rev.58:491-562)。特别地,甲病毒结构蛋白(例如由辅助病毒提供)是用于以蛋白质颗粒形式递送RNA的合适载体。

[0417] 在一个实施方案中,本发明的颗粒制剂是纳米粒制剂。在该实施方案中,根据本发明的组合物包含纳米粒形式的根据本发明的核酸。纳米粒制剂可通过多种方案和用多种络合化合物获得。脂质、聚合物、低聚物或两亲性物质是纳米微粒制剂的典型成分。

[0418] 如本文中所使用的,术语“纳米粒”是指具有使颗粒适合于全身,特别是肠胃外施用的直径的任何颗粒,特别是直径通常为1000纳米(nm)或更小的核酸。在一个实施方案中,纳米粒的平均直径为约50nm至约1000nm,优选约50nm至约400nm,优选约100nm至约300nm,例如约150nm至约200nm。在一个实施方案中,纳米粒的直径为约200至约700nm,约200至约600nm,优选约250至约550nm,特别是约300至约500nm或约200至约400nm。

[0419] 在一个实施方案中,如通过动态光散射测量的本文中所述的纳米粒的多分散指数(polydispersity index,PI)为0.5或更小,优选0.4或更小,或者甚至更优选0.3或更小。

“多分散指数”(PI)是颗粒混合物中单个颗粒(例如脂质体)的均匀或不均匀尺寸分布的量度,并指示混合物中颗粒分布的宽度。例如,可如WO 2013/143555 A1中所述来确定PI。

[0420] 如本文中使用的,术语“纳米粒制剂”或类似术语是指包含至少一个纳米粒的任何颗粒制剂。在一些实施方案中,纳米粒组合物是纳米粒的均匀集合。在一些实施方案中,纳米粒组合物是含有脂质的药物制剂,例如脂质体制剂或乳剂。

[0421] 含有脂质的药物组合物

[0422] 在一个实施方案中,本发明的药物组合物包含至少一种脂质。优选地,至少一种脂质是阳离子脂质。所述含有脂质的药物组合物包含根据本发明的核酸。在一个实施方案中,根据本发明的药物组合物包含包封在囊泡(例如脂质体)中的RNA。在一个实施方案中,根据本发明的药物组合物包含乳剂形式的RNA。在一个实施方案中,根据本发明的药物组合物包含与阳离子化合物复合的rRNA,从而形成例如所谓的脂质复合物或多聚复合物。RNA包封在囊泡(例如脂质体)中不同于例如脂质/RNA复合物。脂质/RNA复合物可例如在RNA例如与预先形成的脂质体混合时获得。

[0423] 在一个实施方案中,根据本发明的药物组合物包含包封在囊泡中的rRNA。这样的制剂是根据本发明的特定的颗粒制剂。囊泡是卷成球形壳的脂双层,其封闭小的空间并且将该空间与囊泡外部的空间分隔开。通常而言,囊泡内部的空间是水性空间,即包含水。通常而言,囊泡外部的空间是水性空间,即包含水。脂双层由一种或更多种脂质(形成囊泡的脂质)形成。包围囊泡的膜是层状相,类似于质膜的层状相。根据本发明的囊泡可以是多层囊泡、单层囊泡或其混合物。当包封在囊泡中时,rRNA通常与任何外部介质分离。因此,其以受保护的形式存在,功能上等同于天然甲病毒中的受保护形式。合适的囊泡是如本文中所述的颗粒,特别是纳米粒。

[0424] 例如,RNA(rRNA)可被包封在脂质体中。在该实施方案中,药物组合物是或包含脂质体制剂。脂质体中的包封通常会保护RNA免受RNase消化。脂质体可包含一些外部RNA(例如在其表面上),但是至少一半的RNA(并且理想地是其全部)包封在脂质体的核心内。

[0425] 脂质体是微观的脂质囊泡,通常具有形成囊泡的脂质(例如磷脂)的一个或更多个双层,并且能够包封药物,例如RNA。在本发明的背景下可使用不同类型的脂质体,包括但不限于多层囊泡(multilamellar vesicle,MLV)、小单层囊泡(small unilamellar vesicle,SUV)、大单层囊泡(large unilamellar vesicle,LUV)、空间稳定脂质体(sterically stabilized liposome,SSL)、多囊囊泡(multivesicular vesicle,MV)和大多囊囊泡(large multivesicular vesicle,LMV)以及本领域已知的其他双层形式。脂质体的大小和层状性(lamellarity)将取决于制备方式。存在超分子组织的多种其他形式,其中脂质可存在于水性介质中,包括由单层构成的层状相、六角相和反六角相、立方相、胶束、反胶束。这些相也可与DNA或RNA组合获得,并且与RNA和DNA的相互作用可能会显著影响相态。这样的相可存在于本发明的纳米粒RNA制剂中。

[0426] 可使用技术人员已知的标准方法形成脂质体。相应方法包括反向蒸发法、乙醇注入法、脱水-再水化法、超声处理或其他合适的方法。在脂质体形成之后,可对脂质体进行大小调整以获得具有基本均一大小范围的脂质体群体。

[0427] 在本发明的一个优选实施方案中,rRNA存在于包含至少一种阳离子脂质的脂质体中。相应脂质体可由单一脂质或由脂质混合物形成,前提是使用至少一种阳离子脂质。优选

的阳离子脂质具有能够被质子化的氮原子;优选地,这样的阳离子脂质是具有叔胺基的脂质。具有叔胺基的特别合适的脂质是1,2-二亚油基氧基-N,N-二甲基-3-氨基丙烷(DLinDMA)。在一个实施方案中,根据本发明的RNA存在于如WO 2012/006378 A1中所述的脂质体制剂中:具有脂双层的脂质体,所述脂双层包封包含RNA的水性核心,其中脂双层包含pKa为5.0至7.6的脂质,其优选具有叔胺基。具有叔胺基的优选的阳离子脂质包括DLinDMA (pKa 5.8),并且一般在WO 2012/031046 A2中描述。根据WO 2012/031046 A2,包含相应化合物的脂质体特别适合于RNA的包封,并因此适合RNA的脂质体递送。在一个实施方案中,根据本发明的RNA存在于脂质体制剂中,其中脂质体包含至少一种阳离子脂质,所述阳离子脂质的头基包含至少一个能够被质子化的氮原子(N),其中脂质体和RNA的N:P比率为1:1至20:1。根据本发明,“N:P比率”是指在包含脂质的颗粒(例如脂质体)中包含的阳离子脂质中的氮原子(N)与RNA中的磷酸酯原子(P)的摩尔比,如WO 2013/006825A1中所述。1:1至20:1的N:P比率与脂质体的净电荷和RNA向脊椎动物细胞的递送效率相关。

[0428] 在一个实施方案中,根据本发明的rRNA存在于脂质体制剂中,所述脂质体制剂包含至少一种脂质,所述脂质包含聚乙二醇(PEG)部分,其中RNA被包封在PEG化脂质体中,使得PEG部分存在于脂质体外部,如WO 2012/031043 A1和WO 2013/033563 A1中所述。

[0429] 在一个实施方案中,根据本发明的rRNA存在于脂质体制剂中,其中脂质体的直径为60至180nm,如WO 2012/030901 A1中所述。

[0430] 在一个实施方案中,根据本发明的rRNA存在于脂质体制剂中,其中含rRNA的脂质体具有接近于零或为负的净电荷,如WO 2013/143555 A1中所公开。

[0431] 在另一些实施方案中,根据本发明的rRNA以乳剂的形式存在。先前已经描述了用于将核酸分子(例如rRNA分子)递送至细胞的乳剂。本文中优选的是水包油乳剂。相应乳剂颗粒包含油核心和阳离子脂质。更优选阳离子水包油乳剂,其中根据本发明的RNA与乳剂颗粒复合。乳剂颗粒包含油核心和阳离子脂质。阳离子脂质可与带负电荷的rRNA相互作用,从而将rRNA锚定至乳剂颗粒。在水包油乳剂中,乳剂颗粒分散在水连续相中。例如,乳剂颗粒的平均直径通常可为约80nm至180nm。在一个实施方案中,本发明的药物组合物是阳离子水包油乳剂,其中乳剂颗粒包含油核心和阳离子脂质,如WO 2012/006380 A2中所述。根据本发明的rRNA可以以包含阳离子脂质的乳剂形式存在,其中乳剂的N:P比率为至少4:1,如WO 2013/006834 A1中所述。根据本发明的rRNA可以以阳离子脂质乳剂的形式存在,如WO 2013/006837 A1中所述。特别地,组合物可包含与阳离子水包油乳剂颗粒复合的rRNA,其中油/脂质的比率为至少约8:1(摩尔:摩尔)。

[0432] 在另一些实施方案中,根据本发明的药物组合物包含脂质复合物形式的RNA。术语“脂质复合物”或“RNA脂质复合物”是指脂质与核酸(例如RNA)的复合物。脂质复合物可由阳离子(带正电)脂质体和阴离子(带负电)核酸形成。阳离子脂质体还可包括中性的“辅助”脂质。在最简单的情况下,通过以一定的混合方案将核酸与脂质体混合而自发形成脂质复合物,但是可应用多种其他方案。应当理解,带正电荷的脂质体和带负电荷的核酸之间的静电相互作用是脂质复合物形成的驱动力(WO 2013/143555A1)。在本发明的一个实施方案中, RNA脂质复合物颗粒的净电荷接近零或为负。已知RNA和脂质体的电中性或带负电荷的脂质复合物在全身施用之后导致脾树突状细胞(dendritic cell, DC)中大量RNA表达,并且与对于带正电荷的脂质体和脂质复合物已报道的升高的毒性无关(参见WO 2013/143555A1)。因

此,在本发明的一个实施方案中,根据本发明的药物组合物包含纳米粒形式的RNA,优选脂质复合物纳米粒形式的RNA,其中(i)纳米粒中的正电荷数量不超过纳米粒中的负电荷数量,以及/或者(ii)纳米粒具有中性或净负电荷,以及/或者(iii)纳米粒中正电荷与负电荷的电荷比为1.4:1或更小,以及/或者(iv)纳米粒的 ζ 电势为0或更小。如WO 2013/143555 A1中所述, ζ 电势是胶体系统中电动势的科学术语。在本发明中,可如WO 2013/143555 A1中公开的那样计算纳米粒中的(a) ζ 电势和(b)阳离子脂质与RNA的电荷比二者。总之,如WO 2013/143555A1中公开的其中颗粒的净电荷接近零或为负的,具有限定粒径的纳米粒脂质复合物制剂的药物组合物是本发明的背景下优选的药物组合物。

[0433] 在一个实施方案中,本文中所述的核酸例如rRNA以脂质纳米粒(LNP)的形式施用。LNP可包含能够形成与一个或多个核酸分子连接或者其中包封一个或多个核酸分子的颗粒的任何脂质。

[0434] 在一个实施方案中,LNP包含一种或更多种阳离子脂质以及一种或更多种稳定脂质。稳定脂质包括中性脂质和聚乙二醇化脂质。

[0435] 在一个实施方案中,LNP包含阳离子脂质、中性脂质、类固醇、聚合物缀合脂质;以及包封在脂质纳米粒内或与脂质纳米粒缔合的RNA。

[0436] 在一个实施方案中,LNP包含40至55mol百分比、40至50mol百分比、41至49mol百分比、41至48mol百分比、42至48mol百分比、43至48mol百分比、44至48mol百分比、45至48mol百分比、46至48mol百分比、47至48mol百分比或47.2至47.8mol百分比的阳离子脂质。在一个实施方案中,LNP包含约47.0、47.1、47.2、47.3、47.4、47.5、47.6、47.7、47.8、47.9或48.0mol百分比的阳离子脂质。

[0437] 在一个实施方案中,中性脂质以5至15mol百分比、7至13mol百分比或9至11mol百分比的浓度范围存在。在一个实施方案中,中性脂质以约9.5、10或10.5mol百分比的浓度存在。

[0438] 在一个实施方案中,类固醇以30至50mol百分比、35至45mol百分比或38至43mol百分比的浓度范围存在。在一个实施方案中,类固醇以约40、41、42、43、44、45或46mol百分比的浓度存在。

[0439] 在一个实施方案中,LNP包含1至10mol百分比、1至5mol百分比或1至2.5mol百分比的聚合物缀合的脂质。

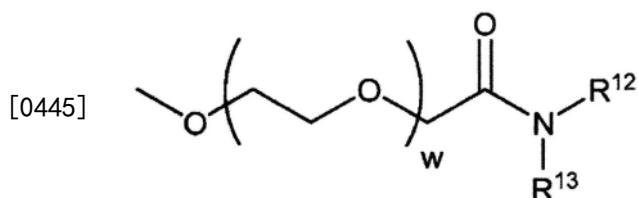
[0440] 在一个实施方案中,LNP包含40至50mol百分比的阳离子脂质;5至15mol百分比的中性脂质;35至45mol百分比的类固醇;1至10mol百分比的聚合物缀合的脂质;以及包封在脂质纳米粒内或与脂质纳米粒缔合的RNA。

[0441] 在一个实施方案中,基于脂质纳米粒中存在的脂质的总mol来确定mol百分比。

[0442] 在一个实施方案中,中性脂质选自DSPC、DPPC、DMPC、DOPC、POPC、DOPE、DOPG、DPPG、POPE、DPPE、DMPE、DSPE和SM。在一个实施方案中,中性脂质选自DSPC、DPPC、DMPC、DOPC、POPC、DOPE和SM。在一个实施方案中,中性脂质是DSPC。

[0443] 在一个实施方案中,类固醇是胆固醇。

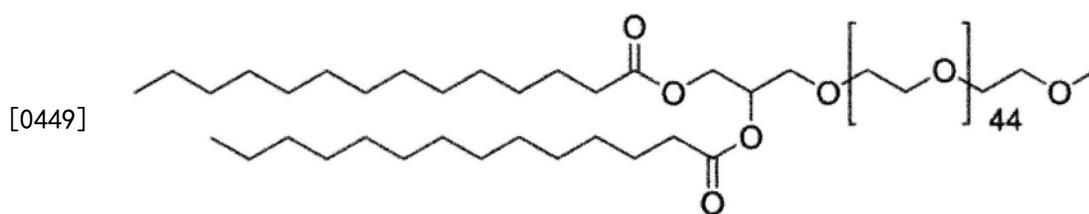
[0444] 在一个实施方案中,聚合物缀合的脂质是聚乙二醇化的脂质。在一个实施方案中,聚乙二醇化脂质具有以下结构。



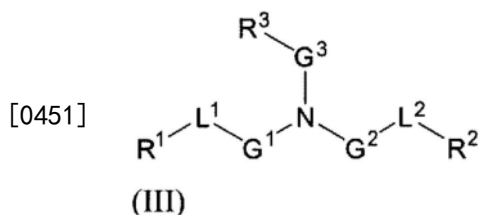
[0446] 或者为其可药用盐、互变异构体或立体异构体,其中:

[0447] R^{12} 和 R^{13} 各自独立地是包含10至30个碳原子的直链或支链、饱和或不饱和烷基链,其中所述烷基链任选地被一个或更多个酯键中断;并且 w 平均值范围为30至60。在一个实施方案中, R^{12} 和 R^{13} 各自独立地是包含12至16个碳原子的直链饱和烷基链。在一个实施方案中, w 平均值范围为40至55。在一个实施方案中,平均 w 为约45。在一个实施方案中, R^{12} 和 R^{13} 各自独立地是包含约14个碳原子的直链饱和烷基链,并且 w 平均值为约45。

[0448] 在一个实施方案中,聚乙二醇化脂质是DMG-PEG 2000,例如具有以下结构:



[0450] 在一些实施方案中,LNP的阳离子脂质组分具有式(III)的结构:



[0452] 或者为其可药用盐、互变异构体、前药或立体异构体,其中:

[0453] L^1 或 L^2 中的一者是 $-O(C=O)-$, $-(C=O)O-$, $-C(=O)-$, $-O-$, $-S(O)_x-$, $-S-S-$, $-C(=O)S-$, $SC(=O)-$, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$ 或 $-NR^aC(=O)O-$,并且 L^1 或 L^2 中的另一者是 $-O(C=O)-$, $-(C=O)O-$, $-C(=O)-$, $-O-$, $-S(O)_x-$, $-S-S-$, $-C(=O)S-$, $SC(=O)-$, $-NR^aC(=O)-$, $-C(=O)NR^a-$, $NR^aC(=O)NR^a-$, $-OC(=O)NR^a-$ 或 $-NR^aC(=O)O-$ 或直接键;

[0454] G^1 和 G^2 各自独立地是未经取代的 C_1-C_{12} 亚烷基或 C_1-C_{12} 亚烯基;

[0455] G^3 是 C_1-C_{24} 亚烷基、 C_1-C_{24} 亚烯基、 C_3-C_8 亚环烷基、 C_3-C_8 亚环烯基;

[0456] R^a 是H或 C_1-C_{12} 烷基;

[0457] R^1 和 R^2 各自独立地是 C_6-C_{24} 烷基或 C_6-C_{24} 烯基;

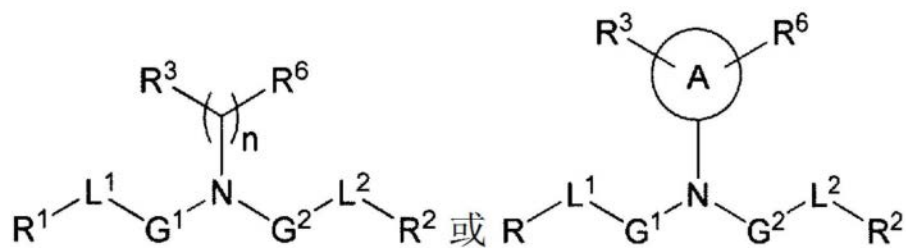
[0458] R^3 是H, OR^5 , CN, $-C(=O)OR^4$, $-OC(=O)R^4$ 或 $-NR^5C(=O)R^4$;

[0459] R^4 是 C_1-C_{12} 烷基;

[0460] R^5 是H或 C_1-C_6 烷基;并且

[0461] x 是0、1或2。

[0462] 在式(III)的一些前述实施方案中,脂质具有以下结构(IIIA)或(IIIB)之一:



(IIIA)

(IIIB)

[0463] 其中:

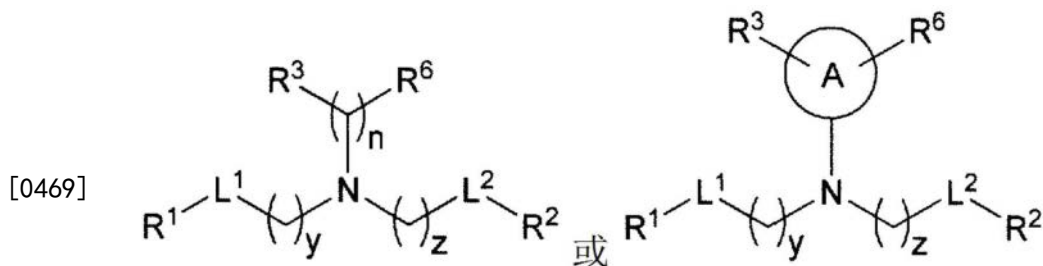
[0464] A是3至8元环烷基或亚环烷基环；

[0465] R^6 在每次出现时独立地是H、OH或 C_1 - C_{24} 烷基;

[0466] n是1至15的整数。

[0467] 在式(III)的一些前述实施方案中,脂质具有结构(IIIA),而在另一些实施方案中,脂质具有结构(IIIB)。

[0468] 在式(III)的另一些实施方案中,脂质具有以下结构(IIIC)或(IIID)之一:



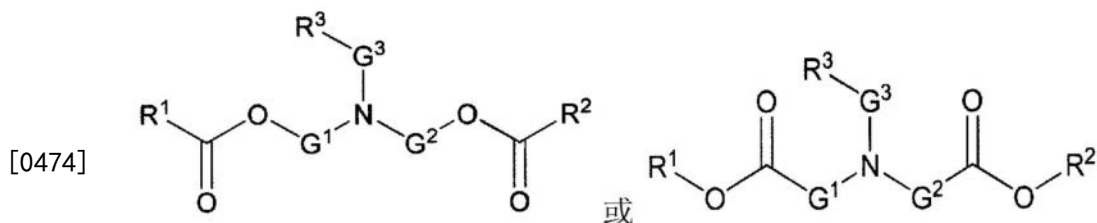
(IIIC)

(IID)

[0470] 其中y和z各自独立地是1至12的整数。

[0471] 在式(III)的任一前述实施方案中, L^1 或 L^2 中的一者是 $-O(C=O)-$ 。例如, 在一些实施方案中, L^1 和 L^2 中的每一者都是 $-O(C=O)-$ 。在任意前述中的一些不同实施方案中, L^1 和 L^2 各自独立地是 $-(C=O)O-$ 或 $-O(C=O)-$ 。例如, 在一些实施方案中, L^1 和 L^2 中的每一者都是 $-(C=O)O-$ 。

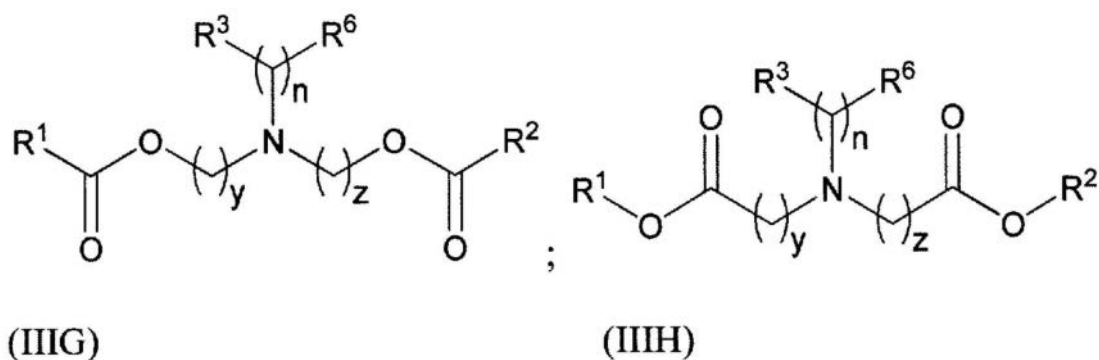
[0472] 在式(III)的一些不同实施方案中,脂质具有以下结构(IIIE)或(IIIF)之一:



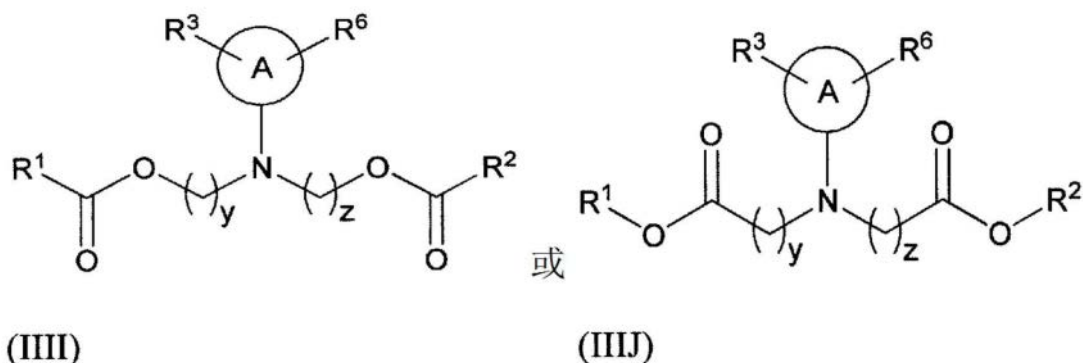
(IIIE)

(IIIF)

[0475] 在式(III)的一些前述实施方案中,脂质具有以下结构(IIIG)、(IIIH)、(IIII)或(IIIJ)之一:



[0476]



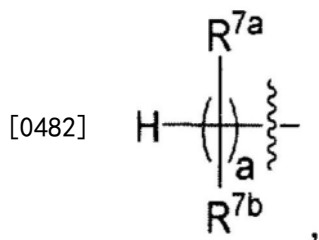
[0477] 在式(III)的一些前述实施方案中, n 是2至12的整数,例如2至8或2至4。例如,在一些实施方案中, n 是3、4、5或6。在一些实施方案中, n 是3。在一些实施方案中, n 是4。在一些实施方案中, n 是5。在一些实施方案中, n 是6。

[0478] 在式(III)的另一些前述实施方案中, y 和 z 各自独立地是2至10的整数。例如,在一些实施方案中, y 和 z 各自独立地是4至9或4至6的整数。

[0479] 在式(III)的一些前述实施方案中, R^6 是H。在另一些前述实施方案中, R^6 是 C_1 - C_{24} 烷基。在另一些实施方案中, R^6 是OH。

[0480] 在式(III)的一些实施方案中, G^3 是未经取代的。在另一些实施方案中, G^3 是经取代的。在多个不同的实施方案中, G^3 是直链 C_1 - C_{24} 亚烷基或直链 C_1 - C_{24} 亚烯基。

[0481] 在式(III)的另一些前述实施方案中, R^1 或 R^2 或二者都是 C_6 - C_{24} 烯基。例如,在一些实施方案中, R^1 和 R^2 各自独立地具有以下结构:



[0483] 其中:

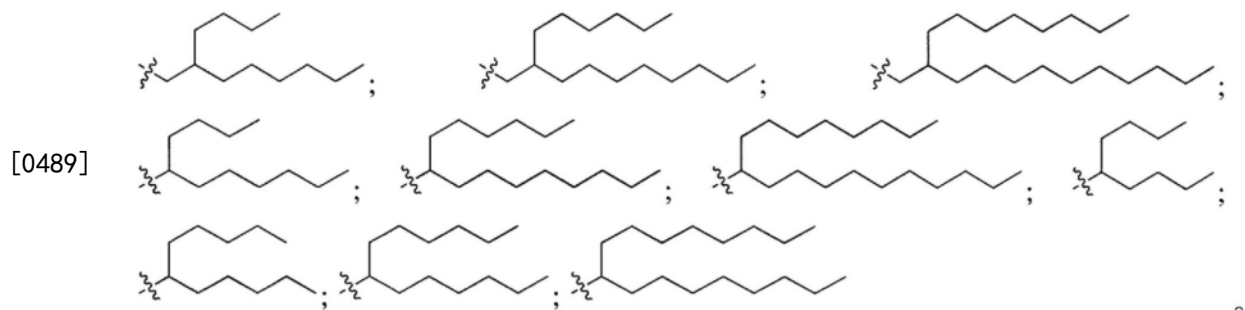
[0484] R^{7a} 和 R^{7b} 在每次出现时独立地是H或 C_1 - C_{12} 烷基;并且

[0485] a 是2至12的整数,

[0486] 其中 R^{7a} 、 R^{7b} 和 a 各自被选择使得 R^1 和 R^2 各自独立地包含6至20个碳原子。例如,在一些实施方案中, a 是5至9或8至12的整数。

[0487] 在式(III)的一些前述实施方案中, R^{7a} 的至少一次出现是H。例如,在一些实施方案中, R^{7a} 在每次出现时都是H。在前述的另一些不同实施方案中, R^{7b} 的至少一次出现是 C_1-C_8 烷基。例如,在一些实施方案中, C_1-C_8 烷基是甲基、乙基、正丙基、异丙基、正丁基、异丁基、叔丁基、正己基或正辛基。

[0488] 在式(III)的不同实施方案中, R^1 或 R^2 或二者具有以下结构之一:



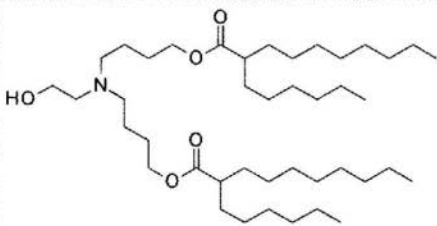
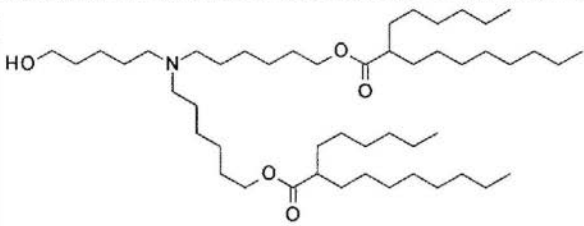
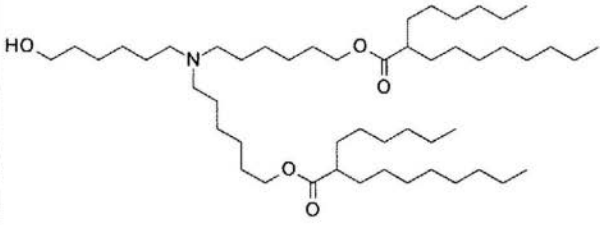
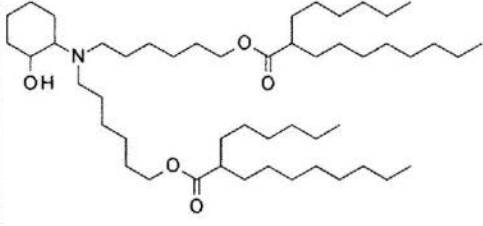
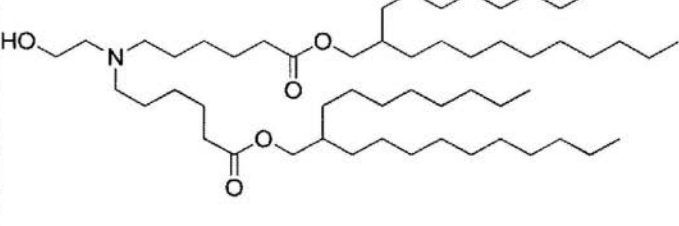
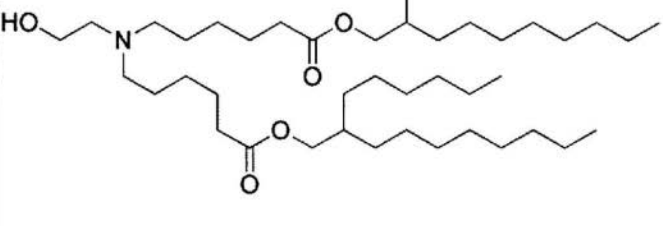
[0490] 在式(III)的一些前述实施方案中, R^3 是OH、CN、 $-C(=O)OR^4$ 、 $-OC(=O)R^4$ 或 $-NHC(=O)R^4$ 。在一些实施方案中, R^4 是甲基或乙基。

[0491] 在多个不同的实施方案中,式(III)的阳离子脂质具有下表中列出的结构之一。

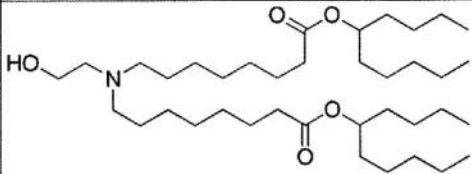
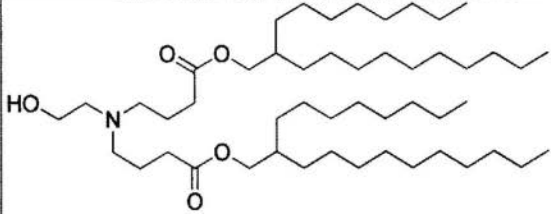
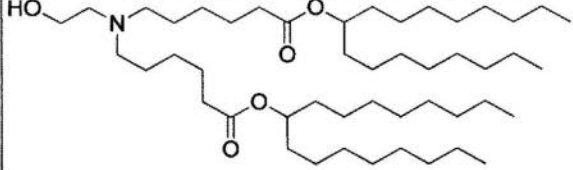
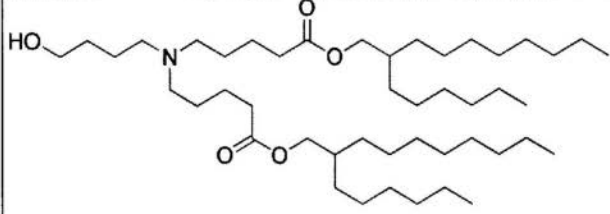
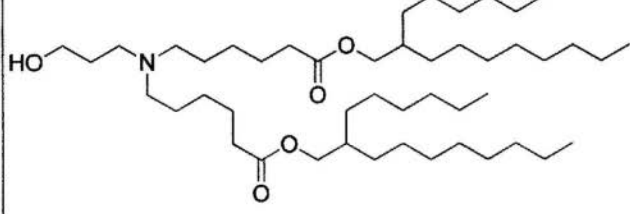
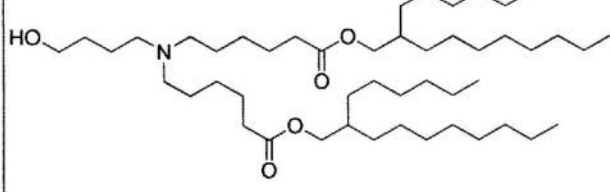
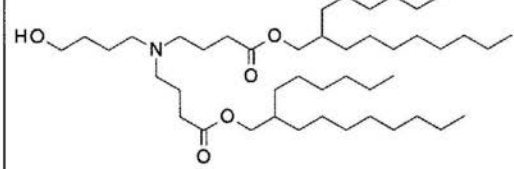
[0492] 式(III)的代表性化合物。

No.	结构
III-1	
III-2	
III-3	
III-4	
III-5	

[0493]

No.	结构
III-6	
III-7	
III-8	
III-9	
III-10	
III-11	

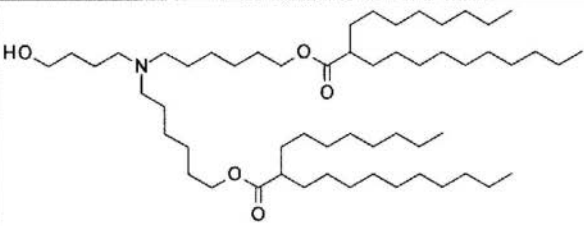
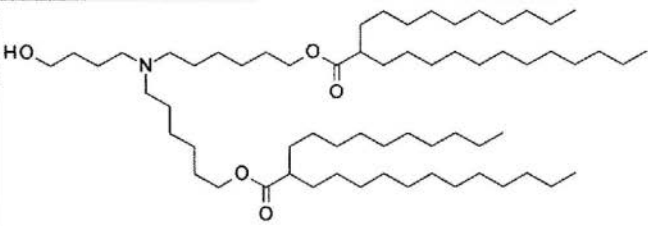
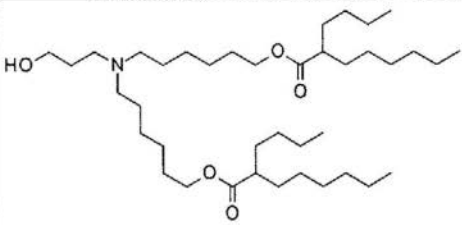
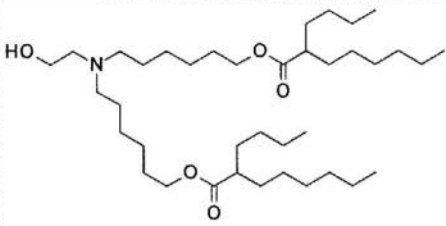
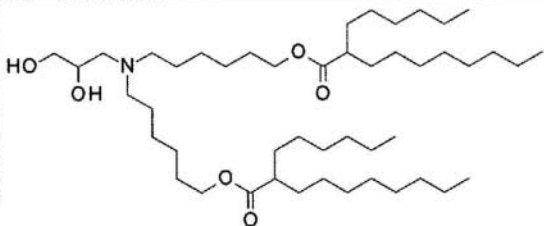
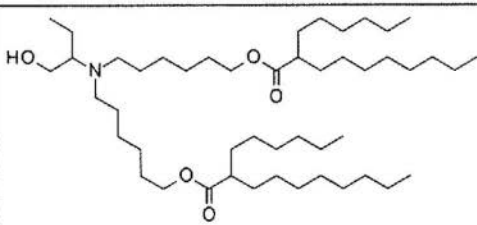
[0494]

No.	结构
III-12	
III-13	
III-14	
III-15	
III-16	
III-17	
III-18	

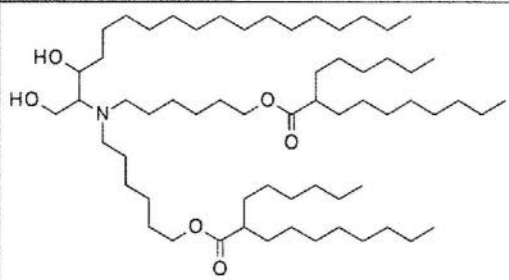
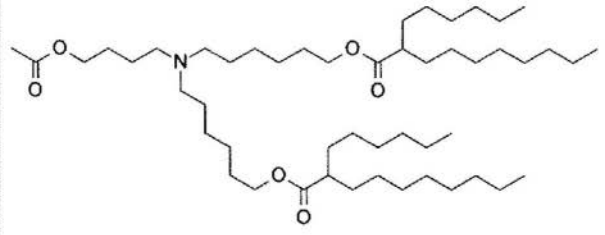
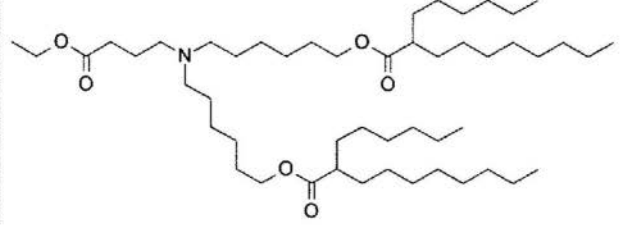
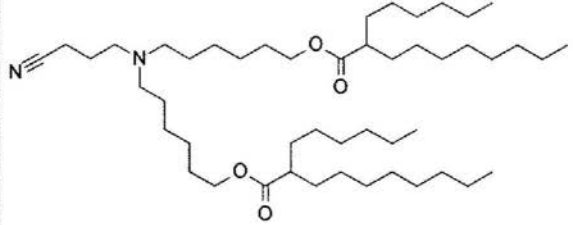
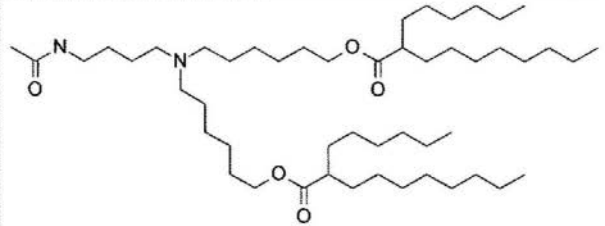
[0495]

No.	结构
III-19	
III-20	
III-21	
III-22	
III-23	
III-24	
III-25	

[0496]

No.	结构
III-26	
III-27	
III-28	
III-29	
III-30	
III-31	

[0497]

No.	结构
III-32	
III-33	
III-34	
III-35	
III-36	

[0498]

[0499] 在一些实施方案中,LNP包含式(III)的脂质、RNA、中性脂质、类固醇和聚乙二醇化脂质。在一些实施方案中,式(III)的脂质是化合物III-3。在一些实施方案中,中性脂质是DSPC。在一些实施方案中,类固醇是胆固醇。在一些实施方案中,聚乙二醇化脂质是ALC-0159。

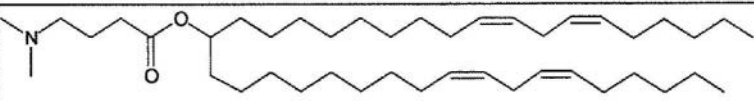
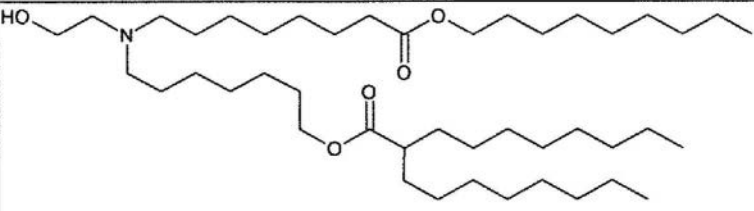
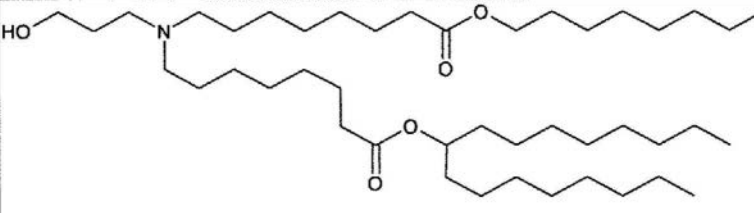
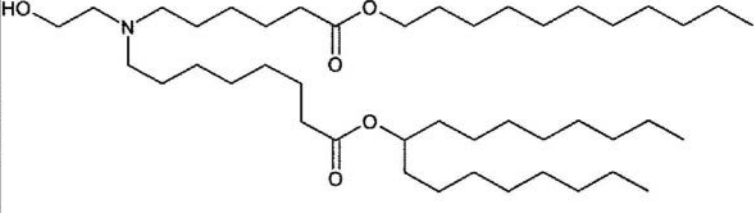
[0500] 在一些实施方案中,阳离子脂质以约40至约50摩尔百分比的量存在于LNP中。在一

个实施方案中,中性脂质以约5至约15摩尔百分比的量存在于LNP中。在一个实施方案中,类固醇以约35至约45摩尔百分比的量存在于LNP中。在一个实施方案中,聚乙二醇化脂质以约1至约10摩尔百分比的量存在于LNP中。

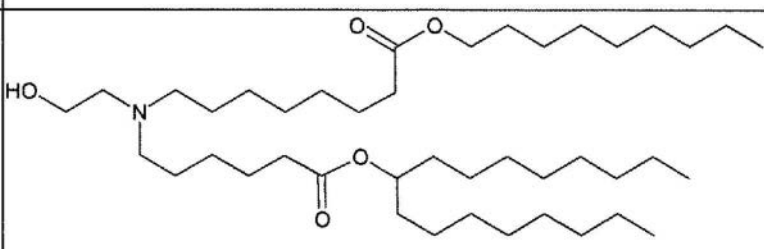
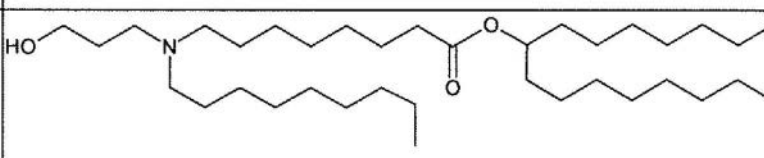
[0501] 在一些实施方案中,LNP以约40至约50摩尔百分比的量包含化合物III-3、以约5至约15摩尔百分比的量包含DSPC、以约35至约45摩尔百分比的量包含胆固醇和以约1至约10摩尔百分比的量包含ALC-0159。

[0502] 在一些实施方案中,LNP以约47.5摩尔百分比的量包含化合物III-3、以约10摩尔百分比的量包含DSPC、以约40.7摩尔百分比的量包含胆固醇和以约1.8摩尔百分比的量包含ALC-0159。

[0503] 在多个不同的实施方案中,阳离子脂质具有下表中列出的结构之一。

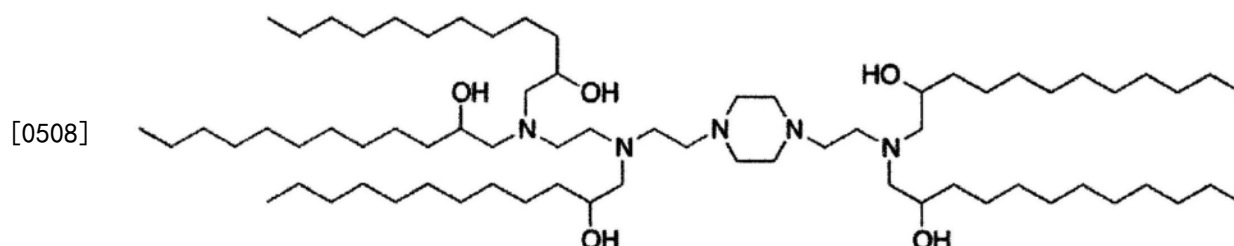
No.	结构
A	
B	
C	
D	

[0504]

No.	结构
E	
F	

[0506] 在一些实施方案中,LNP包含上表中示出的阳离子脂质,例如式(B)或式(D)的阳离子脂质,特别是式(D)的阳离子脂质、RNA、中性脂质、类固醇和聚乙二醇化脂质。在一些实施方案中,中性脂质是DSPC。在一些实施方案中,类固醇是胆固醇。在一些实施方案中,聚乙二醇化脂质是DMG-PEG 2000。

[0507] 在一个实施方案中,LNP包含为可电离的类脂质物质(类脂质(lipidoid))的阳离子脂质。在一个实施方案中,阳离子脂质具有以下结构:



[0509] N/P值优选为至少约4。在一些实施方案中,N/P值范围为4至20、4至12、4至10、4至8或5至7。在一个实施方案中,N/P值为约6。

[0510] 本文中所述LNP可具有在一个实施方案中范围为约30nm至约200nm,或约60nm至约120nm的平均直径。

[0511] RNA靶向

[0512] 本公开内容的一些方面涉及本文中所公开的rRNA(例如,编码疫苗抗原和/或免疫刺激剂的RNA)的靶向递送。

[0513] 在一个实施方案中,本公开内容涉及靶向肺。如果施用的RNA是编码疫苗抗原的RNA,则特别优选靶向肺。RNA可被递送至肺,例如通过吸入施用可被配制为如本文中所述颗粒(例如,脂质颗粒)的RNA。

[0514] 在一个实施方案中,本公开内容涉及靶向淋巴系统,特别是次级淋巴器官,更特别是脾。如果施用的RNA是编码疫苗抗原的RNA,则特别优选靶向淋巴系统,特别是次级淋巴器官,更特别是脾。

[0515] 在一个实施方案中,靶细胞是脾细胞。在一个实施方案中,靶细胞是抗原呈递细胞,例如脾中的专职抗原呈递细胞。在一个实施方案中,靶细胞是脾中的树突状细胞。

[0516] “淋巴系统”是循环系统的一部分并且是免疫系统的重要部分,其包含运送淋巴的淋巴管网络。淋巴系统由淋巴器官、淋巴管的传导网络和循环淋巴组成。初级或中央淋巴器官从未成熟的祖细胞产生淋巴细胞。胸腺和骨髓构成初级淋巴器官。次级或外周淋巴器官(包括淋巴结和脾)维持成熟的初始淋巴细胞并启动适应性免疫应答。

[0517] RNA可通过所谓的脂质复合物制剂递送至脾,其中RNA与包含阳离子脂质和任选地另外的脂质或辅助脂质的脂质体结合以形成可注射的纳米粒制剂。脂质体可通过将脂质在乙醇中的溶液注射到水或合适的水相中而获得。RNA脂质复合物颗粒可通过将脂质体与RNA混合来制备。靶向脾的RNA脂质复合物颗粒在WO 2013/143683中描述,其通过引用并入本文。已发现具有净负电荷的RNA脂质复合物颗粒可用于优先靶向脾组织或脾细胞,例如抗原呈递细胞,特别是树突状细胞。因此,在施用RNA脂质复合物颗粒之后,在脾中发生RNA积聚和/或RNA表达。因此,本公开内容的RNA脂质复合物颗粒可用于在脾中表达RNA。在一个实施方案中,在施用RNA脂质复合物颗粒之后,在肺和/或肝中没有发生或基本上没有发生RNA积聚和/或RNA表达。在一个实施方案中,在施用RNA脂质复合物颗粒之后,在脾中的抗原呈递细胞例如专职抗原呈递细胞中发生RNA积聚和/或RNA表达。因此,本公开内容的RNA脂质复合物颗粒可用于在这样的抗原呈递细胞中表达RNA。在一个实施方案中,抗原呈递细胞是树突状细胞和/或巨噬细胞。

[0518] 本公开内容的RNA脂质复合物颗粒的电荷是至少一种阳离子脂质中存在的电荷与RNA中存在的电荷的总和。电荷比是至少一种阳离子脂质中存在的正电荷与RNA中存在的负电荷的比。至少一种阳离子脂质中存在的正电荷与RNA中存在的负电荷的电荷比通过以下等式来计算:电荷比=[(阳离子脂质浓度(mol))* (阳离子脂质中的正电荷总数)]/[(RNA浓度(mol))* (RNA中的负电荷总数)]。

[0519] 本文中所述的靶向脾的RNA脂质复合物颗粒在生理pH下优选地具有净负电荷,例如正电荷与负电荷的电荷比为约1.9:2至约1:2、或约1.6:2至约1:2、或约1.6:2至约1.1:2.0。在一些具体实施方案中,在生理pH下RNA脂质复合物颗粒中正电荷与负电荷的电荷比为约1.9:2.0、约1.8:2.0、约1.7:2.0、约1.6:2.0、约1.5:2.0、约1.4:2.0、约1.3:2.0、约1.2:2.0、约1.1:2.0或约1:2.0。

[0520] 免疫刺激剂可通过向对象施用编码免疫刺激剂的RNA而提供给对象,所述RNA在用于将RNA优先递送至肝或肝组织的制剂中。优选将RNA递送至这样的靶器官或组织,特别是如果期望表达大量的免疫刺激剂和/或如果期望或需要免疫刺激剂全身性存在,特别是以显著量存在。

[0521] RNA递送系统对肝具有固有的偏好。这涉及基于脂质的颗粒、阳离子和中性纳米粒,特别是脂质纳米粒,例如脂质体、纳米胶束和亲脂性配体,其在生物缀合物中。肝积聚是由肝脉管系统或脂质代谢(脂质体和脂质或胆固醇缀合物)的不连续性质引起的。

[0522] 为了在体内将RNA递送至肝,可使用药物递送系统通过阻止RNA降解将RNA转运到肝中。例如,由聚(乙二醇)(PEG)包被的表面和含有mRNA的核组成的复合物(polyplex)纳米胶束是可用的系统,因为纳米胶束在生理条件下在体内提供了优异的RNA稳定性。此外,由密集的PEG栅栏构成的复合物纳米胶束表面所提供的隐形(stealth)特性有效地规避了宿主免疫防御。

[0523] 用于靶向肝的合适免疫刺激剂的一些实例是参与T细胞增殖和/或维持的细胞因

子。合适的细胞因子的一些实例包括IL2或IL7、其片段和变体,以及这些细胞因子、片段和变体的融合蛋白,例如PK延长的细胞因子。

[0524] 在另一个实施方案中,编码免疫刺激剂的RNA可以在用于优先将RNA递送至淋巴系统,特别是次级淋巴器官,更特别是脾的制剂中施用。优选将免疫刺激剂递送至这样的靶组织,特别是如果期望在该器官或组织中存在免疫刺激剂(例如,用于诱导免疫应答,特别是在T细胞致敏期间或用于活化驻留免疫细胞而需要免疫刺激剂例如细胞因子的情况下),而不期望免疫刺激剂全身性地存在,特别是以显著量存在(例如,由于免疫刺激剂具有全身性毒性)。

[0525] 合适的免疫刺激剂的一些实例是参与T细胞致敏的细胞因子。合适的细胞因子的一些实例包括IL12、IL15、IFN- α 或IFN- β 、其片段和变体,以及这些细胞因子、片段和变体的融合蛋白,例如PK延长的细胞因子。

[0526] 用于产生蛋白质的方法

[0527] 本发明还提供了用于在细胞中生产目的蛋白质的方法,其包括以下步骤:

[0528] (a) 获得根据本发明的RNA复制子,其包含编码目的蛋白质的开放阅读框,以及

[0529] (b) 将RNA复制子接种到细胞中。

[0530] 在该方法的多个实施方案中,RNA复制子如上文针对本发明的RNA复制子所限定,只要RNA复制子包含编码目的蛋白质的开放阅读框,任选地编码功能性非结构蛋白的开放阅读框并且可由功能性非结构蛋白复制即可,其中rRNA可包含至少一个经修饰的核苷酸和调节序列中的一个或更多个点突变,所述点突变恢复或提高经修饰rRNA的功能。

[0531] 可接种一个或更多个核酸分子的细胞可称为“宿主细胞”。根据本发明,术语“宿主细胞”是指可用外源核酸分子转化或转染的任何细胞。术语“细胞”优选是完整的细胞,即具有完整膜的细胞,其未释放其正常的胞内组分,例如酶、细胞器或遗传物质。完整的细胞优选是可生存细胞,即能够执行其正常代谢功能的活细胞。术语“宿主细胞”根据本发明包括原核细胞(例如大肠杆菌(E.coli))或真核细胞(例如人和动物细胞、植物细胞、酵母细胞和昆虫细胞)。特别优选的是哺乳动物细胞,例如来自人、小鼠、仓鼠、猪、驯养动物(包括马、牛、绵羊和山羊)以及灵长类的细胞。细胞可来源于多种组织类型并包含原代细胞和细胞系。一些具体实例包括角质形成细胞、外周血白细胞、骨髓干细胞和胚胎干细胞。在另一些实施方案中,宿主细胞是抗原呈递细胞,特别是树突状细胞、单核细胞或巨噬细胞。核酸可以以单拷贝或多拷贝存在于宿主细胞中,并且在一个实施方案中在宿主细胞中表达。

[0532] 细胞可以是原核细胞或真核细胞。原核细胞在本文中是合适的,例如用于根据本发明的DNA的增殖,以及真核细胞在本文中是合适的,例如用于复制子的开放阅读框的表达。

[0533] 在本发明的方法中,可使用根据本发明的任何RNA复制子或根据本发明的试剂盒或根据本发明的药物组合物。RNA可以以药物组合物的形式使用,或者作为裸RNA使用,例如用于电穿孔。

[0534] 在根据本发明的用于在细胞中产生蛋白质的方法中,细胞可以是抗原呈递细胞,并且该方法可用于表达编码抗原的RNA。为此,本发明可涉及将编码抗原的RNA引入到抗原呈递细胞(例如树突状细胞)中。为了转染抗原呈递细胞(例如树突状细胞),可使用包含编码抗原的RNA的药物组合物。

[0535] 在一个实施方案中,用于在细胞中产生蛋白质的方法是体外方法。在一个实施方案中,用于在细胞中产生蛋白质的方法不包括通过手术或治疗从人或动物对象中移出细胞。

[0536] 在该实施方案中,可根据本发明接种的细胞向对象施用以在对象中产生蛋白质并向对象提供蛋白质。细胞可相对于对象是自体的、同基因的、同种异体的或异源的。

[0537] 在另一些实施方案中,在用于在细胞中产生蛋白质的方法中的细胞可存在于对象例如患者中。在这些实施方案中,用于在细胞中产生蛋白质的方法是体内方法,其包括向对象施用RNA分子。

[0538] 在该方面,本发明还提供了用于在对象中产生目的蛋白质的方法,其包括以下步骤:

[0539] (a) 获得根据本发明的RNA复制子,其包含编码目的蛋白质的开放阅读框,以及

[0540] (b) 向对象施用RNA复制子。

[0541] 在该方法的多个实施方案中,RNA复制子如上文针对本发明的RNA复制子所限定,只要RNA复制子包含编码目的蛋白质的开放阅读框,任选地编码功能性非结构蛋白的开放阅读框并且可由功能性非结构蛋白复制即可,其中rRNA可包含至少一个经修饰的核苷酸以及调节序列中的一个或更多个点突变,所述点突变恢复或提高经修饰rRNA的功能。

[0542] 可在根据本发明的用于在对象中产生蛋白质的方法中使用任何根据本发明的RNA复制子或根据本发明的试剂盒或根据本发明的药物组合物。例如,在本发明的方法中,RNA可以以例如如本文中所述的药物组合物的形式或作为裸RNA使用。

[0543] 鉴于向对象施用的能力,根据本发明的每种RNA复制子或根据本发明的试剂盒或根据本发明的药物组合物可称为“药物”等。本发明预见到本发明的RNA复制子、试剂盒和药物组合物被提供用作药物。该药物可用于治疗对象。“治疗”意指向对象施用如本文中所述的化合物或组合物或其他实体。该术语包括通过治疗来治疗人体或动物体的方法。

[0544] 上述药物通常不包含DNA,并且因此与现有技术(例如WO 2008/119827A1)中描述的DNA疫苗相比具有另外的安全特征。

[0545] 根据本发明的替代医学用途包括根据本发明的用于在细胞中产生蛋白质的方法,其中细胞可以是抗原呈递细胞,例如树突状细胞,然后将所述细胞引入至对象。例如,可将编码药学活性蛋白质(例如抗原)的RNA离体引入(转染)到抗原呈递细胞中,例如取自对象的抗原呈递细胞中,并且可将任选地离体克隆增殖的抗原呈递细胞重新引入到相同或不同的对象中。可使用本领域已知的任何方法将转染的细胞重新引入到对象中。

[0546] 可根据本发明的药物施用于有此需要的对象。本发明的药物可用于治疗对象的预防以及治疗方法中。

[0547] 以有效量施用根据本发明的药物。“有效量”涉及单独或与其他剂(dose)一起的足以引起反应或期望的作用的量。在治疗对象中的某种疾病或某种病症的情况下,期望的作用是疾病进展的抑制。这包括疾病进展的减速,特别是疾病进展的中断。治疗疾病或病症的期望的作用也可以是疾病爆发的延迟或疾病爆发的抑制。

[0548] 有效量将取决于所治疗的病症,疾病的严重性,患者的个体参数包括年龄、生理状况、体型(size)和体重,治疗持续时间,伴随治疗的类型(如果有的话),具体的施用方式以及其他因素。

[0549] 疫苗接种

[0550] 术语“免疫接种”或“疫苗接种”通常是指出于治疗性或预防性原因而治疗对象的过程。治疗,特别是预防性治疗,优选地是或包括旨在引发或增强对象例如针对一种或更多种抗原的免疫应答的治疗。根据本发明,如果期望通过使用如本文中所述的rRNA诱导或增强免疫应答,则可通过rRNA引发或增强免疫应答。在一个实施方案中,本发明提供预防性治疗,其优选地是或包括对对象进行疫苗接种。本发明的一个实施方案特别适用于疫苗接种,其中复制子编码作为目的蛋白质的药学活性肽或药学活性蛋白质,所述药学活性肽或药学活性蛋白质是免疫活性化合物或抗原。

[0551] 先前已经描述了RNA用于针对包括病原体的外来物质或癌症的疫苗接种(最近由Ulmer et al., 2012, Vaccine 30:4414-4418综述)。与现有技术中的常规方法相比,根据本发明的复制子是用于有效疫苗接种的特别合适的元件,因为其能够被如本文中所述的功能性甲病毒非结构蛋白复制。根据本发明的疫苗接种可用于例如引发对弱免疫原性蛋白质的免疫应答。在根据本发明的RNA疫苗的情况下,蛋白质抗原从未暴露于血清抗体,而是在RNA翻译之后由转染的细胞自身产生。因此,过敏反应不应成为问题。因此,本发明允许患者重复免疫接种而没有变态反应的风险。

[0552] 在涉及根据本发明的疫苗接种的方法中,将本发明的药物施用于对象,特别是如果期望治疗患有涉及抗原的疾病或处于染上涉及抗原的疾病的风险中的对象。

[0553] 在涉及根据本发明的疫苗接种的方法中,由根据本发明的复制子所编码的目的蛋白质编码例如待由免疫应答针对的细菌抗原,待由免疫应答针对的病毒抗原,或待由免疫应答针对的癌症抗原,或待由免疫应答针对的单细胞生物体的抗原。疫苗接种的效力可通过已知的标准方法评估,例如通过测量来自生物体的抗原特异性IgG抗体。在涉及根据本发明的变应原特异性免疫治疗的方法中,由根据本发明的复制子所编码的目的蛋白质编码与变态反应相关的抗原。变应原特异性免疫治疗(也称为脱敏作用)定义为向具有一种或更多种变态反应的生物体施用优选提高剂量的变应原疫苗,以达到与随后暴露于致病变应原的相关的症状减轻的状态。变应原特异性免疫治疗的效力可通过已知的标准方法评估,例如通过测量来自生物体的变应原特异性IgG和IgE抗体。

[0554] 本发明的药物可施用于对象,例如用于治疗对象,包括对象的疫苗接种。

[0555] 术语“对象”涉及脊椎动物,特别是哺乳动物。例如,在本发明的上下文中的哺乳动物是人,非人灵长类,驯养的哺乳动物例如狗、猫、绵羊、牛、山羊、猪、马等,实验动物例如小鼠、大鼠、兔、豚鼠等,以及圈养的动物例如动物园的动物。术语“对象”还涉及非哺乳动物脊椎动物,例如鸟类(特别是家禽,例如鸡、鸭、鹅、火鸡)和鱼类(特别是养殖鱼类,例如鲑鱼或鲟鱼)。如本文中使用的术语“动物”还包括人。

[0556] 在一些实施方案中,向驯养动物例如狗、猫、兔、豚鼠、仓鼠、绵羊、牛、山羊、猪、马、鸡、鸭、鹅、火鸡或野生动物例如狐狸施用是优选的。例如,根据本发明的预防性疫苗接种可适合于向动物群体接种疫苗,例如,在农业中,或野生动物群体。圈养的其他动物群体例如宠物或动物园的动物可接种疫苗。

[0557] 施用方式

[0558] 根据本发明的药物可以以任何合适的途径施加至对象。

[0559] 例如,药物可全身性施用,例如静脉内(intravenously, i.v.)、皮下

(subcutaneously, s.c.)、皮内 (intradermally, i.d.) 或通过吸入。

[0560] 在一个实施方案中,将根据本发明的药物例如皮下施用于肌肉组织,例如骨骼肌或皮肤。通常理解的是,将RNA转移到皮肤或肌肉中导致高度和持续的局部表达,同时强烈诱导体液和细胞免疫应答 (Johansson et al., 2012, PLoS. One. 7: e29732; Geall et al., 2012, Proc. Natl. Acad. Sci. U.S.A 109: 14604-14609)。

[0561] 施用于肌肉组织或皮肤的替代方案包括但不限于:皮内、鼻内、眼内、腹膜内、静脉内、间质、经颊、透皮或舌下施用。皮内和肌内施用是两种优选的途径。

[0562] 可以以多种方式实现施用。在一个实施方案中,通过注射施用根据本发明的药物。在一个优选的实施方案中,通过针头注射。可使用无针注射作为替代方案。

[0563] 通过附图和实施例详细描述和举例说明本发明,这些附图和实施例仅用于举例说明目的并不意味着限制。根据描述和实施例,技术人员可获得同样包括在本发明中的另一些实施方案。

附图说明

[0564] 图1. 载体设计 (未按比例绘制)。(1A) 由甲病毒基因组构建的自扩增RNA (saRNA或rRNA)。saRNA是具有编码甲病毒RNA依赖性RNA聚合酶(复制酶)的5' -开放阅读框(ORF)和编码目的基因的3' -开放阅读框(ORF)的双顺反子RNA。ORF标有AUG起始密码子。saRNA的编码区侧翼是病毒5' 和3' 非翻译区(vUTR)。此外,其包含由RNA复制所需的保守序列元件(CSE)构成的调节区(CSE1和CSE2包含基因组正链启动子(5' -复制识别序列RRS)、CSE4核心基因组负链启动子和CSE3亚基因组启动子)。值得注意的是,5' -和3' -CSE二者协同启动负链合成。复制酶ORF的起始密码子位于5' -调节区内。(1B) 用于定向进化的反式扩增RNA系统。甲病毒复制酶由具有非病毒5' 和3' UTR以及长poly-A尾(pA)的非复制mRNA编码。该mRNA与编码细胞表面表达的CD90.1的反式复制子一起共转染,所述CD90.1用作可选择的表面标志物。构建反式复制子使得5' RRS缺少AUG,并且编码序列从其自身明确的AUG开始插入下游。该构建体允许核苷酸变化而不影响可选择的标志物基因。

[0565] 图2. 反式复制子模板和RNA合成。(2A) 将编码反式复制子的CD90.1插入到质粒中,所述质粒通过将T7-RNA聚合酶启动子连接至5' 末端的正向引物和在终止密码子重叠的CD90.1 ORF的3' 末端处结合的反向引物进行PCR扩增。反向引物将最小的3' CSE和短poly-A (30个腺苷)与3' 末端连接。将所得的PCR产物用作T7聚合酶驱动的RNA体外转录(IVT)的模板,使用常规核苷酸ATP、CTP、GTP和UTP或者ATP、CTP、GTP和N1-甲基-假UTP。(2B) 复制酶mRNA合成。编码T7启动子、5' UTR、复制酶ORF、3' UTR和长poly A尾(100A)的质粒在poly-A尾的末端处用限制酶线性化。使用常规核苷酸将该线性化质粒体外转录为表达复制酶的未经修饰(含有UTP)的mRNA。

[0566] 图3. 体外进化过程。以PCR产物为模板,将编码CD90.1的初始(N=0)反式复制子(TR-CD90.1)体外转录(IVT)为未经修饰的RNA(含有尿苷(U),或经修饰的RNA(含有N1-甲基假尿苷(1mY(或m1ψ))。将每种体外转录的RNA都与编码复制酶的mRNA一起电穿孔到细胞中。第二天,通过流式细胞术评估CD90.1表面表达。最初,包含U的可复制RNA或反式复制子(TR)的表达比包含1mY代替的TR高得多,如在理想化的条形图中所示。经修饰RNA的总表达相比于未经修饰RNA的总表达的比率(1mY指数)远低于1。接下来,将表达CD90.1的经修饰TR所转

染的细胞通过磁辅助细胞分选 (magnetic assisted cell sorting, MACS) 富集, 进行裂解并且使用TR特异性引物通过一步法RT-PCR将细胞TR拷贝转化为cDNA和新的PCR产物。然后再次使用该循环N+1PCR产物来合成U和1mY TR-RNA用于新的细胞电穿孔。将该过程重复N次直至观察到1mY指数随着循环次数的提高而稳健且可重复的提高, 如在理想化的条形图中所示, 并且如本文中所显示的观察到的。

[0567] 图4. TR-CD90.1在11个体外进化循环之后获得了在1mY修饰的情况下提高的相容性。在11个体外进化循环之后, 使用提高量的1mY代替UTP将得到的细胞提取的和逆转录TR进行体外转录。这些经增加修饰的TR与编码复制酶的mRNA一起被电穿孔到K562细胞中。(4A) 第二天通过流式细胞术评估CD90.1表达。(4B和4B-1; 4B-1代表具有统计学分析的重复实验) 1mY指数计算为经修饰TR与未经修饰TR的CD90.1表达的比率。

[0568] 图5. 在11个体外进化循环之后5' CSE区的点突变的积累。在11个体外进化循环之后, 将所得的细胞提取的和逆转录TR进行PCR扩增并通过NGS进行测序。将突变的TR转录物根据其频率进行排序。这里显示的突变是四种最常见的突变, 或存在于四种最常见的转录物中, 并且被选择用于进一步分析。发现所有其他突变的最大频率均低于0.85%并在此处列出。数字等于核苷酸位置, 其中1是第一个体外转录的G。

[0569] 图6. 在5' CSE中鉴定的点突变提高了经修饰的反式复制子的复制和表达。(6A) 将K562细胞用2.6 μ g未经修饰的编码SFV复制酶的mRNA以及0.03 μ g经修饰(1mY)或未经修饰(UTP)的编码CD90.1的VEEV-NTR进行电穿孔。NTR是非突变体(WT)或者具有如标示出的在5' CSE内鉴定的体外进化的突变。(6B) 第二天, 通过流式细胞术评估由NTR导致的CD90.1表达。转染率由直方图中标记条上方的数字表示。(6C和6C-1; 6C-1代表具有统计学分析的重复实验) 基于流式细胞术数据计算1mY指数。该指数量化了经修饰的NTR与未经修饰的NTR的表达的比率。

[0570] 图7. 来源于VEEV-TC-83株的单个5' 核苷酸交换提高了经修饰和未经修饰的反式复制子响应于SFV复制酶的表达。(7A) 将K562细胞用2.6 μ g未经修饰的针对SFV复制酶的mRNA以及0.03 μ g经修饰(1mY)或未经修饰(UTP)的编码CD90.1的VEEV-NTR进行电穿孔。NTR是非突变体(WT)或者具有TC-83特异性5' 突变(从载体的5' 端G计数, 在第4位处将G交换成a)。(7B) 第二天, 通过流式细胞术评估由NTR导致的CD90.1表达。转染率由直方图中标记条上方的数字表示。(7C, 7C-1; 7C-1代表具有统计学分析的重复实验) 基于流式细胞术数据计算1mY指数。该指数量化了经修饰的NTR与未经修饰的NTR的表达的比率。

[0571] 图8. TC-83特异性5' 突变与通过体外进化鉴定的点突变协同, 以进一步提高由SFV复制酶扩增的经修饰的反式复制子的表达。(8A) 将K562细胞用2.6 μ g未经修饰的针对SFV复制酶的mRNA以及0.03 μ g经修饰(1mY)或未经修饰(UTP)的编码CD90.1的VEEV-NTR进行电穿孔。NTR是非突变体(WT), 携带VEEV-TC-83(TC-83)的5' 突变, 或者另外如标示出的在5' CSE内鉴定的体外进化突变。(8B) 第二天, 通过流式细胞术评估由NTR导致的CD90.1表达。转染率由直方图中标记条上方的数字表示。(8C, 8C-1; 8C-1代表具有统计学分析的重复实验) 基于流式细胞术数据计算1mY指数。该指数量化了经修饰的NTR与未经修饰的NTR的表达的比率。

[0572] 图9. 当复制酶mRNA和体外进化的反式复制子RNA二者被修饰时, 反式扩增RNA的表达在人成纤维细胞中极大地提高。(9A) 将原代人包皮成纤维细胞以50,000细胞/孔接种到

12孔板中,并且第二天用4 μ l MessengerMax和1 μ g总RNA/孔(0.9 μ g SFV复制酶mRNA和0.1 μ g反式复制子)进行脂质体转染。SFV复制酶mRNA(REPL)或NTR用于如所示的不同变体中。第二天,通过流式细胞术评估由NTR导致的CD90.1表达。转染率由直方图中标记上方的数字表示。(9B)转染率显示为条形图。缩写:UTP:未经修饰的RNA;1mY:经修饰的RNA;WT:非突变体;-dsRNA:双链RNA的去除;TC-83:G4A-突变体;3xMut=在67、245、248位处的体外进化突变体)。

[0573] 图10.VEEV-TC-83 5' CSE提高了嵌合VEEV/SFV修饰的自扩增RNA的表达。(10A)为了构建编码萤光素酶的嵌合VEEV/SFV saRNA,去除了SFV-saRNA的5' CSE,包括编码复制酶N端的部分。这部分被对应于缺失的复制酶N端的序列所替换,之前是缺失起始密码子的VEEV 5' CSE。第二个类似的构建体携带VEEV-TC-83特异性5' -突变G4A。将两种构建体的RNA使用MessengerMax脂质体转染到(10B) K562细胞或(10C) 原代人包皮成纤维细胞(human foreskin fibroblast, HFF)中。在72小时内测量萤光素酶表达并且从线图计算曲线下的面积(area under the curve, AUC)以估计总表达。(10D, 10E)将K562或HFF细胞在12孔板中用如在图10B和10C中类似(comparable)但编码GFP的VEEV/SFV saRNA进行脂质体转染。在24小时之后收获细胞并且通过流式细胞术评估转染率。

[0574] 图11.图11A示出了SEQ ID NO:1的序列,其是VEEV特立尼达驴株(登录号L01442)的5' 调节区(第45至47位是nsP1的起始密码子)。图11B示出了实施例中所使用的SEQ ID NO:2的序列,其是VEEV特立尼达驴株(登录号L01442)5' 调节区的经修饰版本。例如,在第1位处添加G残基以使用T7 RNA聚合酶进行有效的体外转录。SEQ ID NO:2中的第280至282位是实施例中编码报道基因的起始密码子。在一个实施方案中,该起始密码子还可用于编码的复制酶。

[0575] 图12.图12A至12D示出了经完全修饰的、1mY适应性复制RNA在小鼠中诱导了稳健的免疫应答。对小鼠进行免疫接种以评价100%经修饰的rRNA是否可用于疫苗接种,以及性能是否与未经修饰的rRNA相当。体液免疫通过(12A)ELISA量化HA特异性IgG抗体或(12B)病毒中和滴度(virus neutralization titer, VNT)进行评估。在用对(12C)MHC-I或(12D)MHC-II具有特异性的HA肽刺激脾细胞的情况下,通过IFN- γ ELISpot评估T细胞免疫。缩写:HA:甲型流感病毒的血凝素;UTP:未经修饰的RNA;1mY:经修饰的RNA;WT:非突变体;TC-83:G4A-突变体;3xMut:在第67、245、248位处的体外进化突变体;ELISA:酶联免疫吸附测定;ELISpot:酶联免疫吸附点;IFN- γ :干扰素 γ ;MHC:主要组织相容性复合体。

[0576] 图13.图13A至13C示出了当经历了与EEEV-复制酶一起进行18个体外进化循环以及与VEEV复制酶一起进行的15个循环时,TR-3xmut获得了在1mY修饰的情况下提高的相容性。(13A)将K562细胞用2.6 μ g编码SFV复制酶或VEEV复制酶的mRNA以及54ng经修饰或未经修饰的TR进行电穿孔,所述经修饰或未经修饰的TR是在所标示的体外进化过程的轮次之后从提取的细胞RNA的RT-PCR产物转录的,所述体外进化过程最终导致TR-3xmut的识别。第二天通过流式细胞术评估编码的CD90.1选择标志物的表达,并如之前完成的计算1mY指数。图表示出了三个独立实验的平均值和SD。与图A的过程相当,将K562细胞用1.3 μ g编码EEEV(13B)或VEEV(13C)复制酶的mRNA以及5ng或15ng经修饰和未经修饰的TR进行电穿孔,所述TR在所标示的体外进化循环之后从细胞来源的cDNA获得。这里,使用TR-3xmut作为起始TR开始体外进化。同样,在转染之后一天测定1mY指数。

[0577] 实施例

[0578] 材料和方法

[0579] 以下材料和方法用于实施例中。

[0580] 质粒克隆、体外转录、RNA纯化:使用标准技术克隆质粒。本发明实施例中使用的单独质粒的克隆细节如WO 2017/162266 A1和WO 2017/162265 A1的实施例1中所述。简言之,一种质粒编码插入到非复制mRNA的骨架中的委内瑞拉马脑炎病毒特立尼达驴株(VEEV; GenBank登录号L01442)的复制酶,以及第二种质粒编码表达鼠CD90.1的可复制RNA或反式复制子RNA TR(Δ ATG-RRS Δ SGP)。将在体外进化过程期间鉴定的点突变通过基因合成和亚克隆插入到5' Δ ATG-RRS中。视情况而定,将CD90.1 ORF交换成GFP。命名为“TC-83”的载体版本与其他载体版本的不同之处在于在VEEV的减毒疫苗株TC-83中鉴定的5'末端附近的单核苷酸交换。因此,合成RNA的5'末端处的前4个核苷酸是5' GAUG (特立尼达驴) 或5' -GAUA (TC-83)。

[0581] 对于体外转录,这些质粒通过在poly-A下游限制性消化或通过PCR扩增来线性化,并且用作T7 RNA-聚合酶的模板。RNA合成和纯化如先前所述进行(Holtkamp et al., 2006, Blood 108:4009-4017; Kuhn et al., 2010, Gene Ther. 17:961-971)。双链(double-stranded, ds) RNA耗竭是通过如先前所述的纤维素纯化进行的(Baiersdörfer et al., 2019, Mol Ther Nucleic Acids 15:26-35)。根据需要,UTP由N1-甲基假-UTP以如所示的不同比率交换。

[0582] 纯化的体外转录RNA的质量通过分光光度法评估,并在5200片段分析仪(高级分析)上进行分析。实施例中转染到细胞中的所有RNA都是体外转录的RNA (IVT-RNA)。

[0583] 细胞培养:除非另有说明,否则所有生长培养基、抗生素和其他补充剂均由Life Technologies/Gibco提供。胎牛血清(Fetal calf serum, FCS)购自Sigma-Aldrich。人包皮成纤维细胞从System Bioscience (HFF, 新生儿)获得,并在37°C下在含有15%FCS、1%非必需氨基酸、1mM丙酮酸钠的最小必需培养基(minimum essential media, MEM)中培养。将细胞在37°C下在平衡至5%CO₂的湿润气氛中培养。将K-562细胞(ATCC; CCL-243)在补充有10%FCS的RPMI培养基中培养。

[0584] 将RNA转移到细胞中通过电穿孔或通过脂质介导的转染来进行。对于电穿孔,将RNA重悬于最终体积为62.5 μ l/mm的小池间隙尺寸(cuvette gap size)中并在室温下使用方波电穿孔装置(BTX ECM 830, Harvard Apparatus, Holliston, MA, USA)电穿孔到细胞中。对于不同的细胞类型,应用以下设置:HFF (500V/cm, 1脉冲24毫秒(milliseconds, ms)); K562 (650V/cm, 3脉冲8毫秒并且间隔400毫秒)。使用Lipofectamine Messenger MAX按照制造商的说明(Life Technologies, Darmstadt, Germany)进行RNA脂质体转染。将贴壁细胞以大约20,000个细胞/cm²生长区平板接种,并在24小时之后用总量为250ng/cm²的RNA和1 μ l/cm²的Messenger MAX转染。将RNA物质在无RNA酶的Eppendorf管中混合。

[0585] 萤光素酶测定:为了评估萤光素酶在转染细胞中的表达,将转染细胞平板接种于96孔黑色微孔板(Nunc, Langenselbold, Germany)中。萤火虫萤光素酶的检测是用Bright-Glo萤光素酶测定系统根据制造商的说明进行的。使用微孔板发光读数器Infinite M200 (Tecan Group, Männedorf, Switzerland)测量生物发光。将在给定时间点下确定的萤光素酶活性相对于时间绘图,并且通过梯形法则计算曲线下的面积。使用萤光素酶阴性细胞

以评估背景信号。

[0586] 磁性辅助细胞分选 (MACS): 对于活细胞首先选择表达CD90.1的细胞(死细胞去除试剂盒, Miltenyi Biotec, Bergisch-Gladbach; Germany) 并随后使用抗CD90.1微珠 (Miltenyi Biotec, Bergisch-Gladbach; Germany) 通过MACS按照制造商的说明从经转染的细胞团中分离。

[0587] 流式细胞术 (CD90.1; GFP): 对于流式细胞术, 将细胞使用CD90.1/Thy-1.1-PE克隆OX-7抗体 (BD; Heidelberg; Germany) 进行染色, 或者当表达GFP时使其保持未染色。使用BD FACS Canto II流式细胞仪测量标志物表达。使用配套的FACSDiva软件或FlowJo软件完成数据分析。

[0588] 一步法逆转录聚合酶链反应 (一步法RT-PCR): 将表达CD90.1的细胞使用iScript™RT-qPCR样品制备试剂 (BioRad) 裂解。将裂解物使用具有Platinum™Taq DNA聚合酶 (Invitrogen) 的SuperScript™III一步法RT-PCR系统在没有预先提取RNA的情况下用于一步法RT-PCR。

[0589] 聚合酶链反应 (PCR): 为了产生用于体外转录的DNA模板, 将TR (Δ ATG-RRS Δ SGP) 通过特异性正向引物和特异性反向引物扩增, 所述特异性正向引物包含T7启动子, 所述特异性反向引物与CD90.1的3' 末端结合并引入最小的3' CSE和30个腺苷残基的短poly-A。对于PCR, 根据制造商的说明使用Q5®高保真DNA聚合酶 (NEB)。

[0590] 下一代测序: 为了鉴定在11个体外进化循环之后突变的积累, 通过PCR扩增对应于5' -RRS的cDNA, 并使用MiSeq试剂盒v3 (Illumina) 在Illumina MiSeq上对所得的产物进行测序。

[0591] 实施例1

[0592] 从甲病毒基因组构建的自扩增RNA (saRNA) 的特征在于以下不可或缺的结构域和编码区, 如在图1A中所示。关于编码区, saRNA最常见的形式是具有两个开放阅读框 (ORF) 的双顺反子RNA。5' -ORF编码甲病毒非结构多蛋白 (non-structural polyprotein, nsP), 由具有RNA依赖性RNA转录所必需的所有酶功能的4个亚基组成。nsP逐步成熟并构建成蛋白质复合物 (所谓的复制酶), 所述蛋白质复合物与细胞膜相关联, 在细胞膜上其形成用作RNA扩增的隔室 (compartment) 的小球 (spherule)。第二个ORF位于复制酶ORF的下游并通过亚基因组启动子 (SGP) 与复制酶ORF分开, 所述第二个ORF接收目的基因的遗传信息, 其通常是抗原。一些在甲病毒物种之间保守的RNA结构域 (保守序列元件; CSE) 控制与复制酶的相互作用, 并从而控制RNA复制。在5' 末端处, 2个CSE位于较长的高度结构化区域内 (称为5' -复制识别序列; RRS), 高度结构化区域与复制酶的ORF重叠。该区域比其他更下游CSE (CSE3 = 亚基因组启动子以及CSE4 = 3' 端CSE) 长得多并且更结构化。另外, saRNA是5' -加帽并且3' -聚腺苷酸化的。帽对于复制酶翻译是必需的, 但反之对于复制是非必需的。除了使能够有效翻译复制酶之外, poly-A尾的前大约15个腺苷是3' CSE的一部分。

[0593] 对于体外进化的过程, 我们使用非复制mRNA来表达如WO 2017/162265A1 (图1B) 中所述的复制酶和能够被复制酶反式复制的短的可复制RNA (rRNA) (反式复制子; TR)。该TR的特征在于缺乏起始密码子的5' -RRS (Δ ATG-RRS), 以及不具有如WO 2017/162266 A1中示例性描述的亚基因组启动子 (SGP) (Δ SGP), 其下游是用于可检测的报道表面蛋白 (此处: CD90.1) 的ORF以及在3' 末端处最小3' CSE和30个腺苷的短poly-A。

[0594] 实施例2

[0595] 将克隆到质粒中的CD90.1-TR进行PCR扩增。正向引物将T7-RNA聚合酶启动子连接至紧邻5'-调节区上游的5'末端。反向引物与CD90.1ORF的3'末端互补,并且具有包含最小3' CSE和短poly-A(30个腺苷)的非结合突出端。所得的PCR产物是用于体外转录的模板DNA,所述体外转录使用常规核苷酸ATP、CTP、GTP以及UTP或N1-甲基假-UTP的任一者(图2A)。编码复制酶的mRNA使用具有紧邻非甲病毒5' UTR上游的T7启动子的质粒来合成。将质粒在poly-A尾的末端处用限制酶进行线性化。使用常规核苷酸,该线性化的质粒在体外转录成表达复制酶的未经修饰(含UTP)的mRNA(图2B)。

[0596] 实施例3

[0597] 本文中所述的体外进化旨在选择在体外转录过程中并入经修饰核苷酸(例如N1-甲基假尿苷)时保留复制能力的自复制或反式复制RNA。由于这样的经修饰核苷酸的存在, RNA在很大程度上失去了刺激固有免疫受体的能力,从而降低了对体外转录RNA的固有免疫应答,并因此降低了由于施用RNA引起的副作用。然而,这样的经修饰核苷酸的存在还可导致RNA复制/翻译以产生复制酶和/或编码的目的基因(例如抗原)的能力降低。

[0598] 由于其更长的长度,假设5'-调节区相比于短且结构简单的3' CSE为序列进化提供了更多的选择。此外,5' CSE参与负链合成,这是首次使用经修饰的体外转录RNA作为模板的RNA转录。反式复制子和复制酶-mRNA载体以及上述RNA合成过程适合于定向进化策略,该策略在重复循环过程中特异性靶向5'-调节区(图3)。该过程从跨越编码CD90.1的反式复制子的第一PCR产物开始,该产物在体外转录为未经修饰(U)和经修饰(1mY)的RNA(循环次数N=0)。将每种RNA都与未经修饰的编码复制酶的mRNA一起电穿孔到K562细胞中。第二天,通过流式细胞术评估CD90.1表面表达。最初,TR(U)的表达远高于TR(1mY),经修饰RNA相对于未经修饰RNA的总表达的比率(1mY指数)远低于1。接下来,将用表达CD90.1的经修饰TR转染的细胞通过磁性辅助细胞分选(MACS)富集,进行裂解并且使用TR特异性引物通过一步法RT-PCR将细胞TR-拷贝转化为cDNA和新的PCR产物。然后再次使用该循环N+1PCR产物来合成U和1mY TR-RNA来用于电穿孔到细胞中。通过重复该过程,预期1mY指数将随着每一步的进行而提高。

[0599] 实施例4

[0600] TR-CD 90.1在11个体外进化循环之后获得了在1mY修饰的情况下提高的相容性。进行体外进化过程直至观察到1mY指数的稳健且可重复的提高。在11个循环之后,获得来自表达CD90.1的细胞中的TR-RNA并进行逆转录以产生足够量的PCR产物来用于随后的分析。首先,其旨在从体外进化过程独立地确定TR已经进化为在修饰比率(经修饰核苷酸的百分比)提高之后保持复制的反式复制RNA。使用递增比率(50至100%)的1mY替换UTP,对从11个体外进化循环(N=11)产生的TR以及初始TR模板(N=0)进行体外转录。将不断增加的经修饰TR与编码复制酶的mRNA一起电穿孔到K562细胞中。第二天,显示了初始TR的CD90.1表达比11个循环之后的TR对增加的1mY修饰更敏感,如通过流式细胞术进行评估并通过提高的1mY指数进行可视化的(图4)。由于所有1mY百分比的1mY指数都提高了,因此我们得出结论初始序列确实进化了并获得了1mY并入的更大相容性。

[0601] 实施例5

[0602] 表明了11个体外进化循环之后5' 'CSE区的点突变的积累。在实施例4中已经确

定了11个体外进化循环提高了对1mY修饰的相容性,决定对PCR产物进行测序并分析5' 调节区是否积累了特异性突变。下一代测序(NGS)分析揭示了许多点突变,其用数个读数的百分比的频率显示。在17.4%的所有读数中发现了具有相关突变的4个序列的组,其中克隆频率为5.1%至3.6%。这4个克隆具有双重或三重突变,第一个总是在第67位处从A到C的交换,第二个或第三个是在第244至248位处的核苷酸交换(图5)。对于所有来说共有的突变A67C本身排在第24位处。选择4个最常见的克隆和较不常见但共有的突变用于克隆分析。

[0603] 实施例6

[0604] 在5' CSE内鉴定的点突变提高了经修饰的反式复制子的复制和表达。将鉴定的突变残基插入到亲本TR序列中并进行测试。构建了所有克隆的未经修饰和经修饰RNA,并将其与编码复制酶的mRNA共转染到K562细胞中。第二天,通过流式细胞术评估转染率和1mY指数,其显示出所有克隆都比初始未突变的TR(WT)对1mY修饰更具有抗性(图6)。

[0605] 实施例7

[0606] 来源于VEEV-TC-83株的单个5' 核苷酸交换提高了响应于SFV复制酶的经修饰和未经修饰的反式复制子的表达。反式复制的可复制RNA从甲病毒委内瑞拉马脑炎病毒(VEEV)改造,而编码复制酶的mRNA来源于甲病毒塞姆利基森林病毒(SFV)。对于VEEV,几十年前选择了显示减毒表型的疫苗株,即所谓的TC-83株(Berge et al.,1961,American Journal of Epidemiology 73:209-218)。与VEEV的有毒性的特立尼达驴株相比,TC-83含有数个突变。非常好的特征在于第一个CSE内靠近5' 末端的核苷酸交换。这种突变改变了CSE1的5' 端茎环的二级结构,并且这与VEEV-TC-83基因组RNA的复制提高相关(Kulasegaran-Shylini et al.,2009,Virology 387:211-221)。我们观察到这种5' -突变增强了响应于SFV复制酶的反式复制,并提高了经1mY修饰的可复制RNA的复制。然后当在11个体外进化循环之后在TR中存在TC-83 5' -突变时,观察到1mY指数进一步略有提高(图4B和4C)。令人感兴趣的是,响应于编码VEEV复制酶的mRNA,在并入TC-83突变之后TR的1mY指数保持不变(图4D)。

[0607] 实施例8

[0608] TC-83特异性5' -突变与体外进化产生的点突变协同以进一步提高由SFV复制酶扩增的经修饰的反式复制子的表达。接下来,评估了TC-83特异性5' -突变是否会进一步提高通过体外进化选择的突变TR-克隆的相容性。使用PCR,产生了具有和不具有TC-83突变的克隆TR,作为未经修饰或经修饰的RNA。将K562细胞用编码复制酶的mRNA和相同量的所有克隆进行共电穿孔。再次观察到,单独的TC-83特异性突变提高了经修饰的TR的表达,这通过与未经修饰的TR相比更高的转染率和1mY指数来说明。此外,体外进化的点突变体能够进一步提高转染率和1mY指数(图8)。

[0609] 实施例9

[0610] 当复制酶和反式复制子RNA被修饰并携带体外进化的突变时,反式扩增RNA在人成纤维细胞中的表达极大地增加。用于体外进化的K562细胞能够非常有效地表达自复制和反式复制的RNA分子,这表明这些细胞对RNA具有弱的抗病毒响应。尽管这样的细胞系统允许选择对核苷酸修饰具有抗性的复制RNA,但其不适合在存在由模式识别受体触发的强的细胞固有免疫应答的情况下研究经修饰RNA的表达。因此,使用原代人包皮成纤维细胞来研究体外进化的TR的表达。在实验的第一个分支中,使用未经修饰的SFV复制酶mRNA以及未经修饰或经修饰的TR来脂质体转染成纤维细胞。TR是初始非突变序列(WT),初始序列(TC-83)的

TC-83变体,和10个体外进化循环之后鉴定的三重突变克隆,所述三重突变克隆具有TC-83突变(TC-83+3xMut)或不具有TC-83突变(WT+3xMut)。在实验的第二个分支中,将相同的RNA组合进行了共脂质体转染,不同之处在于编码复制酶的mRNA也被1mY修饰,并进一步纯化以去除双链RNA杂质(Baiersdörfer et al., 2019, Mol Ther Nucleic Acids 15:26-35)。当经修饰的复制酶和经修饰的TR组合时,理论上RNA转染具有最小免疫原性。发现成纤维细胞的最大转染率是在低免疫原性分支中并且在三重突变体的经修饰的TC-83变体的情况下实现的(图9)。这些数据表明,产生了非免疫原性反式复制RNA,该RNA保持复制并实现稳健的转基因表达。

[0611] 实施例10

[0612] VEEV-TC-83 5' CSE提高了嵌合VEEV/SFV经修饰的自扩增RNA的表达。已鉴定5' -调节区在修饰之后使能够RNA复制,其旨在构建相应的自复制RNA。首先,使用VEEV 5' 调节区来构建由SFV复制酶复制的RNA。由于VEEV和SFV saRNA之间5' 端区域的简单结构域替换将最有可能导致无功能载体,因此将VEEV或VEEV-TC83的完整5' -调节区与SFV saRNA融合。更确切地说,将包含CSE1的SFV非翻译区删除,并且将VEEV序列直接融合至SFV复制酶起始密码子的上游。所得的嵌合VEEV/SFV的saRNA具有5' VEEV来源的调节区,但编码SFV复制酶和SFV来源的CSE3 (SGP) 和CSE4 (3' CSE) (图10A)。这些载体在体外转录成未经修饰或经修饰的RNA,所述经修饰的RNA以递增的0%至100%的比率用1mY替换尿苷。观察到增多的修饰降低了嵌合saRNA在K562细胞中的复制能力,除非其含有TC-83特异性5' -突变(图10B)。当脂质体转染到具有干扰素活性的原代成纤维细胞中时,嵌合saRNA的TC-83变体更好地表达并且对修饰具有抗性,而非TC83对应物在100%修饰之后显示出降低地表达(图10C)。通过1mY的存在对复制的抑制相比于K562细胞中不明显,这最可能是由于经修饰的saRNA的固有免疫刺激降低导致的偏差。总之,这些数据表明反式复制RNA中具有提高的对核苷酸修饰抗性的5' -调节区被成功地转置到自复制或顺式复制RNA上。

[0613] 用编码GFP的类似嵌合VEEV/SFV saRNA进行了类似的实验。对K562和HFF细胞进行了脂质体转染,并且第二天评估GFP表达以确定细胞的转染率。具有TRD 5' CSE的VEEV/SFV saRNA的转染率随着1mY并入了增加而降低,而TC-83的5' CSE提供了对1mY修饰的抗性以及总体提高的转染率(图10D, 10E)。用编码萤火虫萤光素酶的构建体确认了这些结果。

[0614] 本文中所引用的文件和研究的引证并不旨在承认任何前述内容是相关的现有技术。关于这些文件内容的所有陈述都是基于申请人可获得的信息,并且不构成对这些文件内容正确性的任何承认。

[0615] 呈现以下描述以使本领域普通技术人员能够制造和使用多个实施方案。对特定装置、技术和应用的描述仅作为实例提供。对本文中所述的实例的多种修改对于本领域普通技术人员而言将是明显的,并且在不脱离多个实施方案的精神和范围的情况下,本文中所限定的一般原理可应用于另一些实例和应用。因此,多个实施方案不旨在限于本文中描述和示出的实例,而是与符合权利要求书的范围相一致。

[0616] 实施例11

[0617] 材料和方法:

[0618] 免疫接种:根据FELASA建议以及按照德国动物福利法案和指令(German animal welfare act and Directive) 2010/63/EU,将6至8周龄的雌性BALB/c小鼠保持在无特定病

原体的条件下。所有工作在参考号G-20-8-043下经Landesuntersuchungsamt Rheinland-Palatinate批准。将小鼠用100ng LNP配制的复制酶-mRNA和编码来自流感病毒A/California/7/2009的HA的TR的混合物作为抗原向胫骨前肌中进行肌内(i.m.)免疫接种。间隔28天对小鼠免疫接种两次。每次注射的RNA组合是78ng带有1mY的SFV复制酶mRNA和22ng分别带有U或1mY的TR。所有的RNA都共转录进行加帽。

[0619] ELISpot分析:在第56天分离脾细胞并使用Mabtech小鼠IFN- γ ELISpotPLUS试剂盒进行ELISpot分析。将脾细胞接种到经预包被的ELISpot板上,并在37°C下在湿润的培养箱中用指定的肽合并物(peptide pool)刺激过夜。

[0620]	MHC-I特异性肽合并物	MHC-II特异性肽合并物
	LYEKVRSQL (SEQ ID NO:3)	STRIYQILAIYSTVA (SEQ ID NO:9)
	CPHAGAKSF (SEQ ID NO:4)	KYVKSTKLRLATGLR (SEQ ID NO:10)
	VEPGDKITF (SEQ ID NO:5)	GKCPKYVKSTKLRLA (SEQ ID NO:11)
	SYINDKGKEV (SEQ ID NO:6)	AIYSTVASSLVLVVS (SEQ ID NO:12)
	FYHKCDNT (SEQ ID NO:7)	EELREQ LSSVSSFER (SEQ ID NO:13)
	STASSWSYI (SEQ ID NO:8)	VKLESTRIYQILAIY (SEQ ID NO:14)

[0621] 使用不相关的肽合并物(仅培养基或伴刀豆球蛋白A)进行对照测量。用生物素缀合的抗IFN- γ 抗体使斑点显现,随后用链霉亲和素-碱性磷酸酶(ALP)和5-溴-4-氯-3-吲哚基-磷酸/硝基蓝四唑(5-bromo-4-chloro-3-indolyl-phosphate/nitroblue tetrazolium,BCIP/NBT)底物孵育。将板使用CTL ImmunoSpot®分析仪扫描并通过ImmunoCapture V6.3软件进行分析。以一式三份进行所有测试并且将斑点计数总结为针对每个一式三份的中位值。

[0622] 抗体分析:对于抗体分析,在第一次免疫接种之后的第56天在异氟烷麻醉下采集血液样品。基于WHO关于流感的实验室诊断和病毒学监测的手册(Manual for the Laboratory Diagnosis and Virological Surveillance of Influenza)(WHO全球流感监测网络)进行病毒中和测试(Virus neutralization test,VNT)以确定血清中的中和抗体。通过将小鼠血清的连续稀释液与100组织培养感染剂量50(tissue culture infection dose,TCID₅₀)的活性Cf4/H1N1流感病毒一起孵育2小时,并随后施加到96孔板中的汇合MDCK细胞单层上来定量HA中和抗体的血清水平。在与MDCK孵育3天之后,将细胞培养上清液收获并与0.5%鸡红细胞(RBC;Lohmann Tierzucht,Cuxhaven,Germany)1:2混合。在圆底96孔板中目视观察到RBC凝集,并将VN血清滴度记录为抑制凝集的最低稀释度的倒数(VNT滴度/50 μ l血清)。

[0623] ELISA:根据供应商的方案(Thermo Fisher Scientific,Germany),使用EZ-Link Sulfor-NHS-LC-生物素化试剂盒对重组Cf4/H1N1-HA蛋白(Life Sciences,Idstein,Germany)进行生物素化。将96孔链霉亲和素板(VWR,Darmstadt,Germany)用生物素化HA蛋白在4°C下包被过夜。在洗涤和封闭之后,通过在37°C下在板上孵育1小时筛选针对HA特异性抗体的血清样品。在孵育并包括测定对照之后,将板与经HRP标记的二抗小鼠IgG抗体一起在37°C下孵育另外45分钟,然后施加TMB单底物(TMB one substrate)(BIOTREND,Cologne,Germany)。监测比色检测,并计算出在450nm至参考波长620nm处的光密度读数(λ

450至620nm)。

[0624] 结果:为了解决使用具有TC-83特异性5' 突变和缩写为3xMut的三种突变(TC-83+3xmut)的体外进化TR是否提高了经修饰复制RNA作为疫苗载体的效力的问题,将甲型流感病毒的血凝素基因插入到TR-TC83+3xMut和亲本TR(WT)中。TR-TC83+3xMut RNA在1mY-TP存在下产生,而针对WT TR mRNA使用了UTP。将每种TR以等摩尔比与带有1mY的SFV复制酶mRNA组合,并在LNP中配制。对于初免加强方案中的i.m.疫苗接种,每次注射使用的每种混合物的总剂量为100ng。在56天之后,收集血液以获得血清来确定体液免疫,并且分离脾细胞来评估细胞介导的免疫。通过ELISA测量的HA特异性抗体的终点滴定显示了,与含有U的WT-TR相比,免疫接种了含有1mY的TR-TC83+3xMut的小鼠外周血中循环的HA特异性IgG的平均浓度更高(图12A)。提高的抗体浓度是通过提高的平均病毒中和滴度(VNT)来反映的(图12B)。如通过对CD8⁺T细胞(图12C)和CD4⁺T细胞(图12C)具有特异性的ELISpot分析所示,T细胞免疫也提高了。

[0625] 总之,数据显示用1mY修饰的TR-TC83+3xMut进行免疫接种使能够使用100%1mY修饰的复制RNA作为疫苗载体,而不会损失诱导免疫原性的能力。

[0626] 实施例12

[0627] 使用SFV复制酶进行了11个体外进化循环过程产生了TR,当用1mY修饰时,该TR被稳健地复制(图4B)。为了评估这种提高的对修饰的相容性是否对SFV复制酶具有特异性,将在体外进化数个循环之后从转染细胞cDNA扩增的经修饰和未经修饰的TR与SFV复制酶或与VEEV复制酶共转染(图13A)。当使用SFV复制酶时,通过连续的1mY指数以及体外进化循环反映了对1mY修饰的连续适应,当使用VEEV复制酶复制相同的进化TR时,我们没有观察到那些。这表明在进化过程期间与使用的复制酶结合的适应是有效的。然而,该体外进化过程也应适用于其他甲病毒物种的复制酶。因此,我们旨在评估使用SFV复制酶鉴定的TR(具有已获得的突变67,245和248,称为3xmut)是否会向着在使用VEEV(委内瑞拉马脑炎病毒)或EEEV(东部马脑炎病毒)复制酶下与1mY更好的相容性进化。因此,使用编码VEEV和EEEV复制酶的mRNA代替SFV复制酶,使TR-3xmut经受所述体外进化的另外的循环。重置循环计数,并且在每轮体外进化之后,我们基于具有未经修饰和经修饰TR的CD90.1选择标志物表达来确定1mY指数。我们观察到使用EEEV复制酶(图13B)或VEEV复制酶(图13C)进行的循环越多,1mY指数越高。因此,不论使用的复制酶,本文中所述的体外进化过程均导致复制适应于1mY的TR。

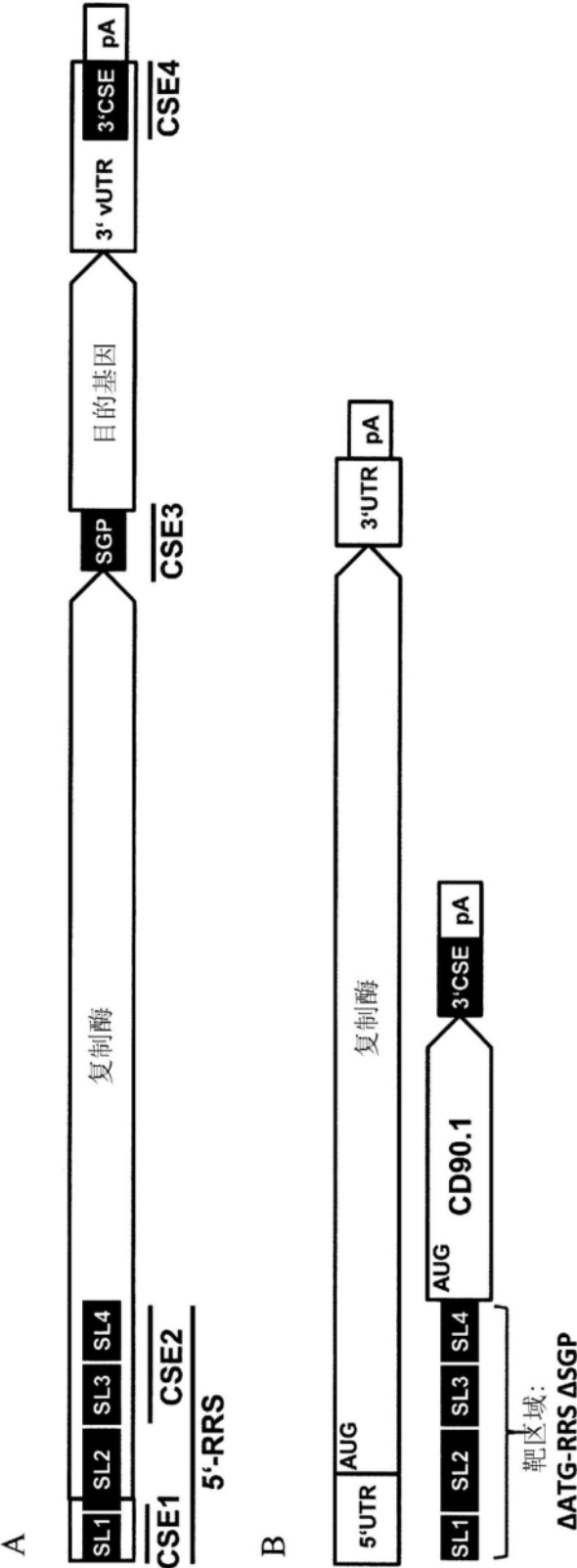


图1A和1B

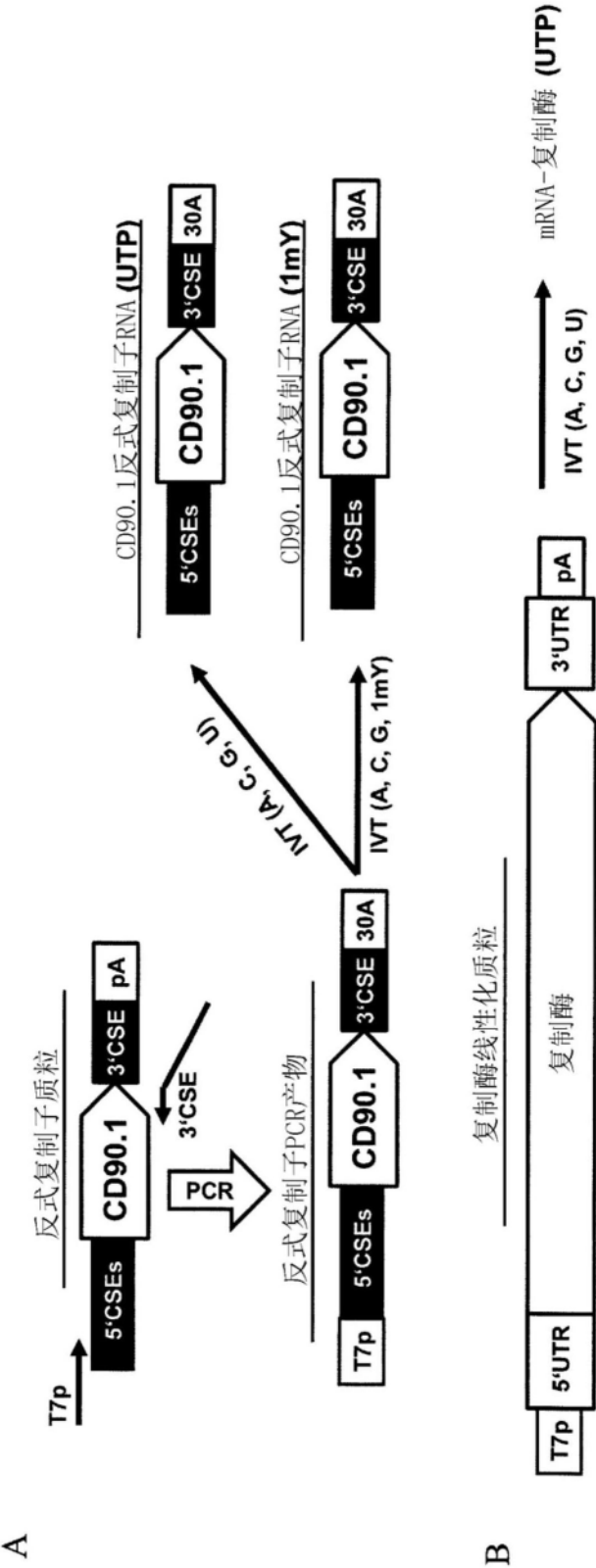


图2A和2B

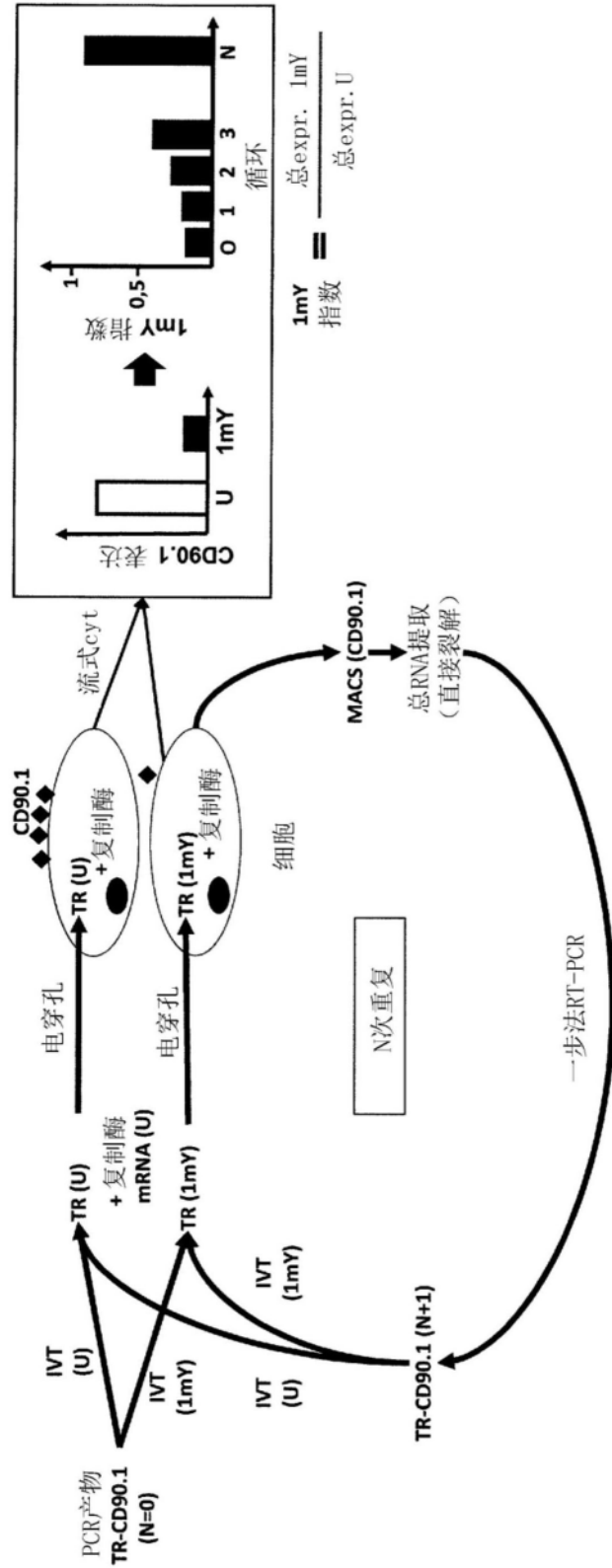


图3

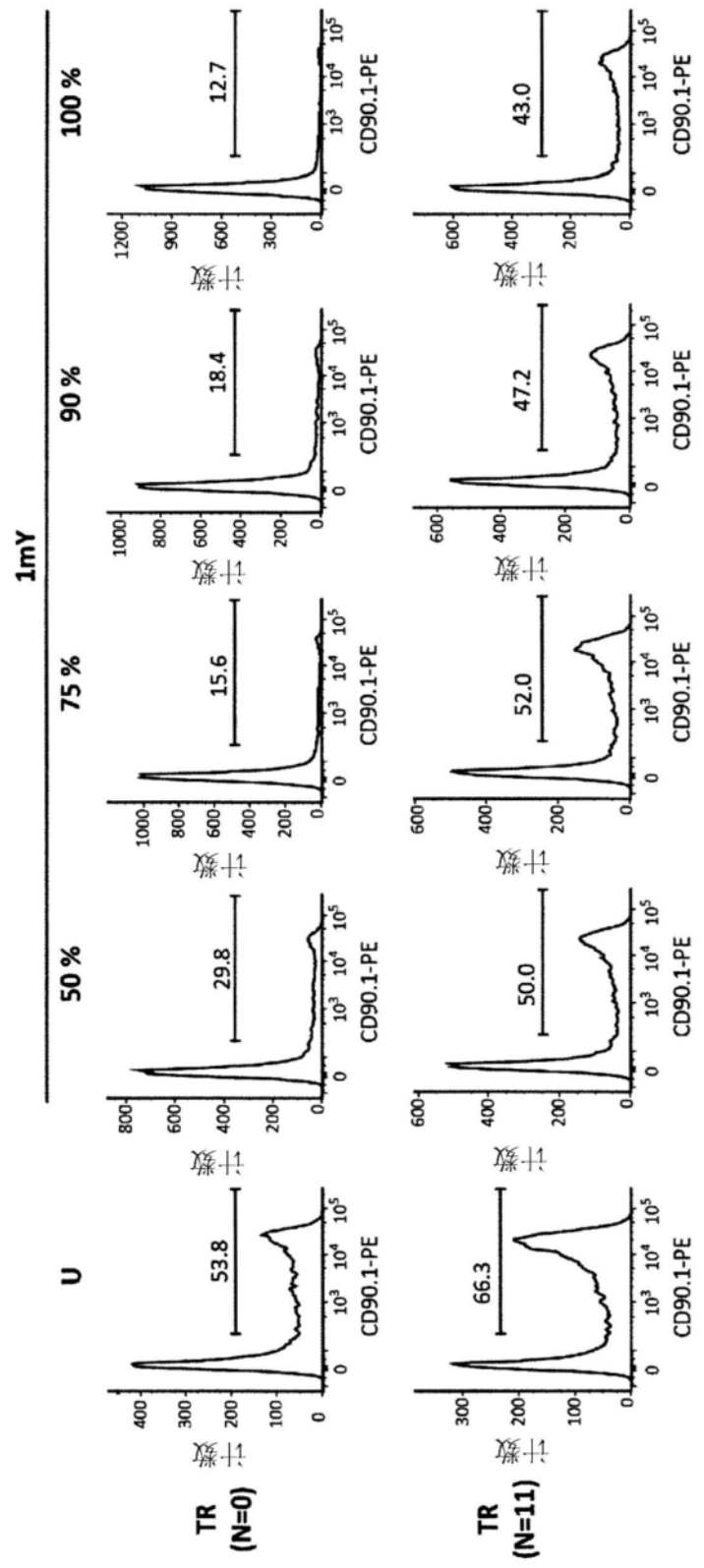


图4A

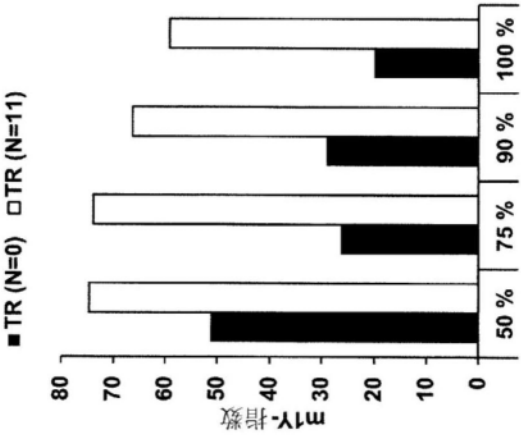


图4B

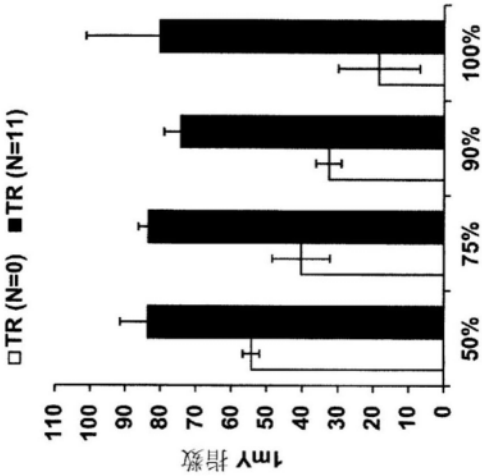


图4B-1

排序	突变	频率[读数的%]
1	A67C + C248A	5.1
2	A67C + G246A	4.8
3	A67C + G244A	3.9
4	A67C + C245A + C248A	3.6
24	A67C	0.1

图5

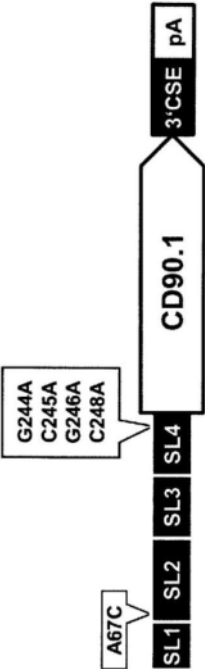


图6A

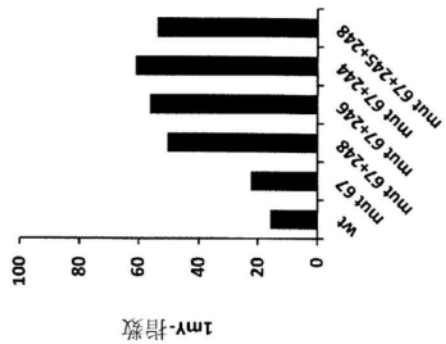


图6C

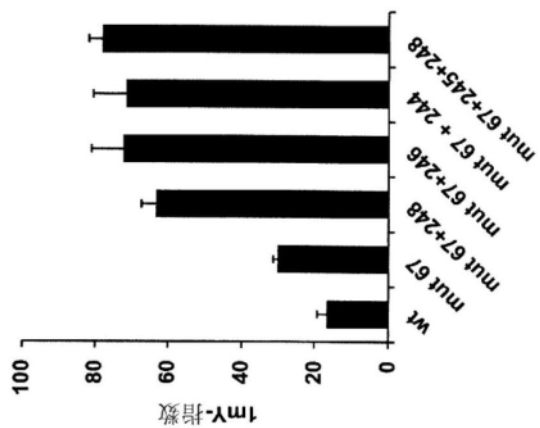


图6C-1

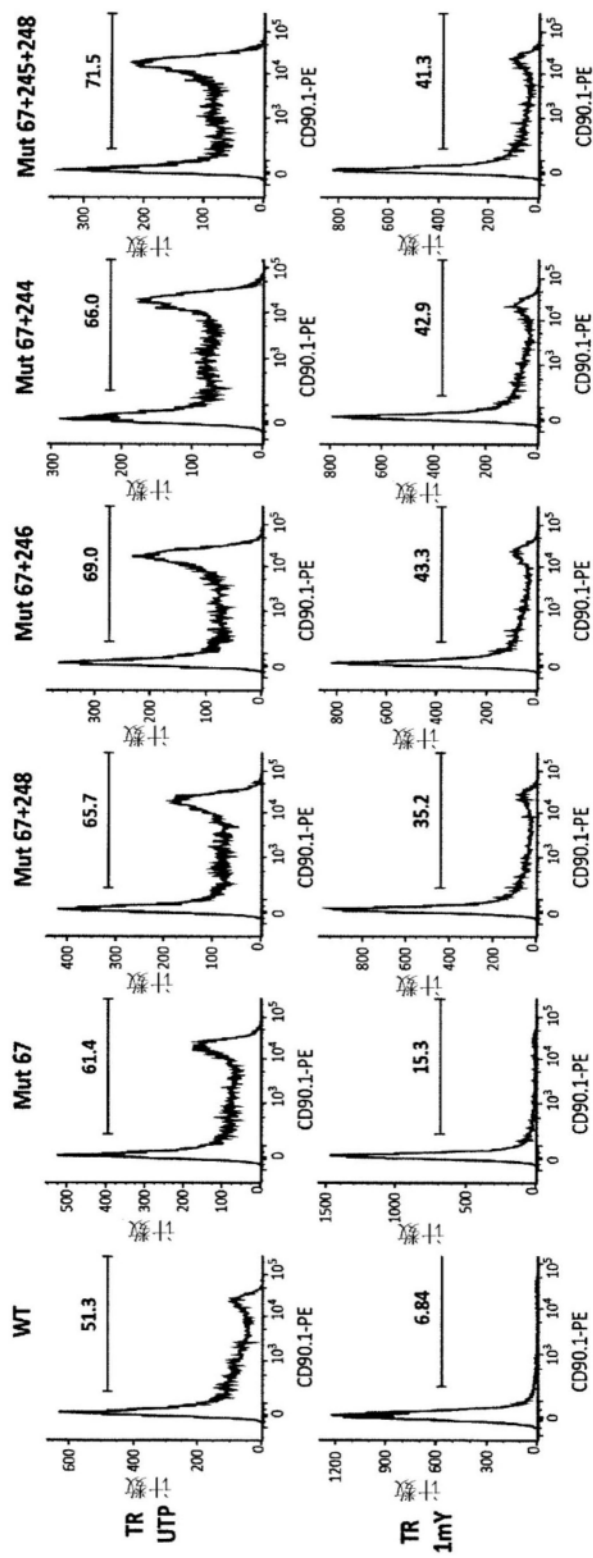


图6B



图7A

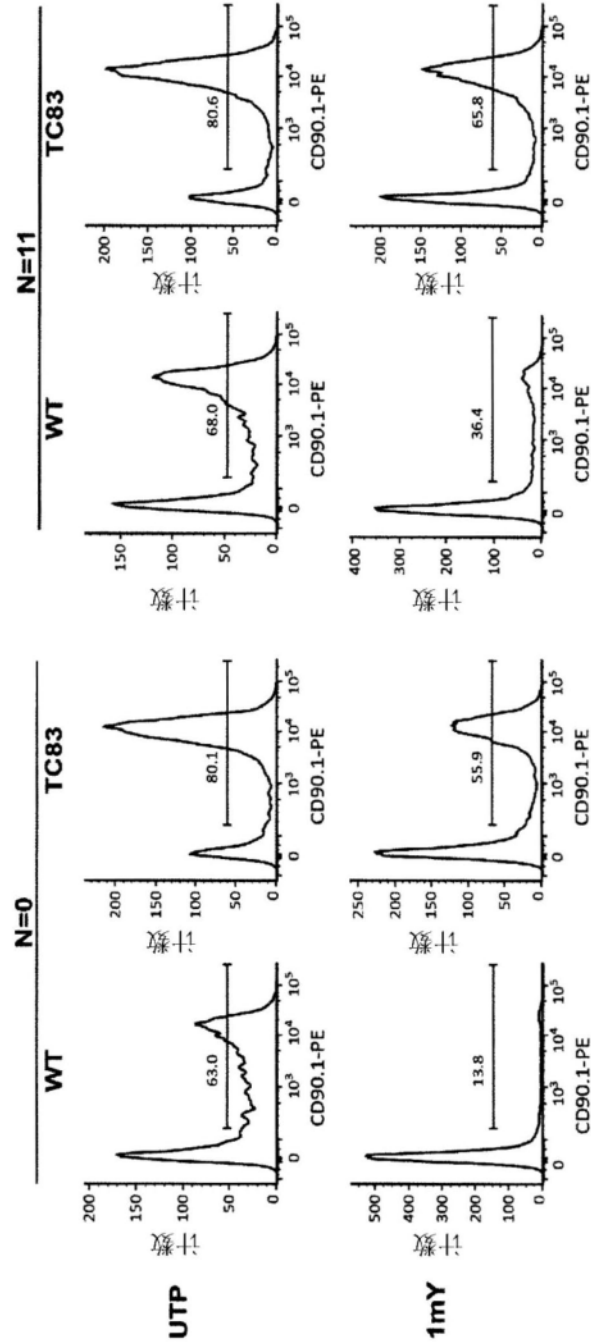


图7B

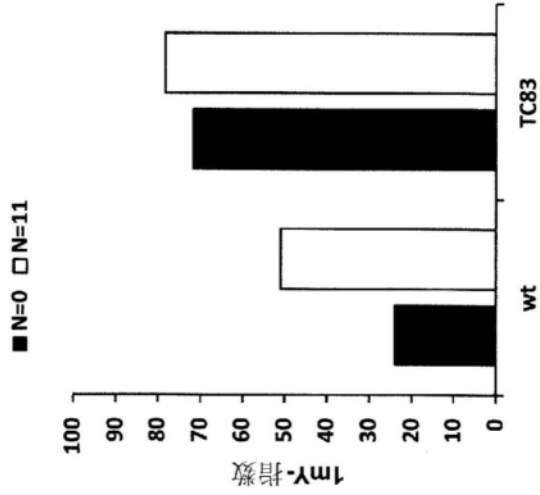


图7C

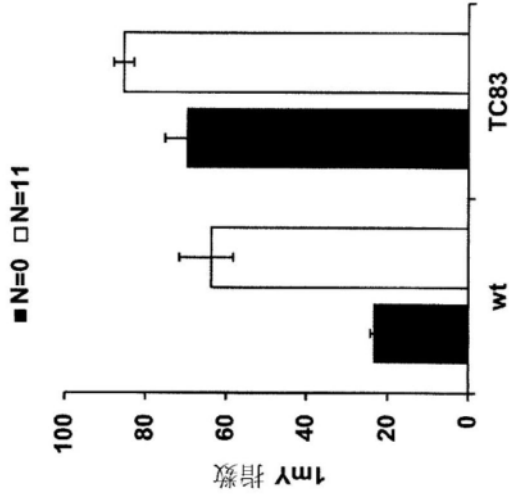


图7C-1

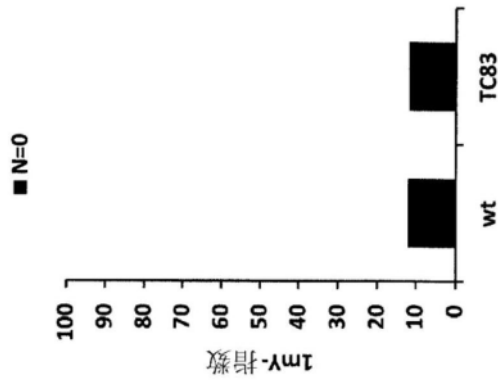


图7D

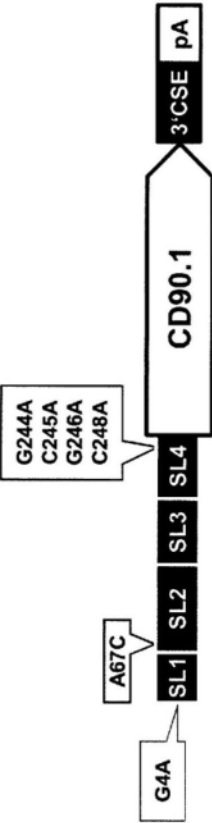


图8A

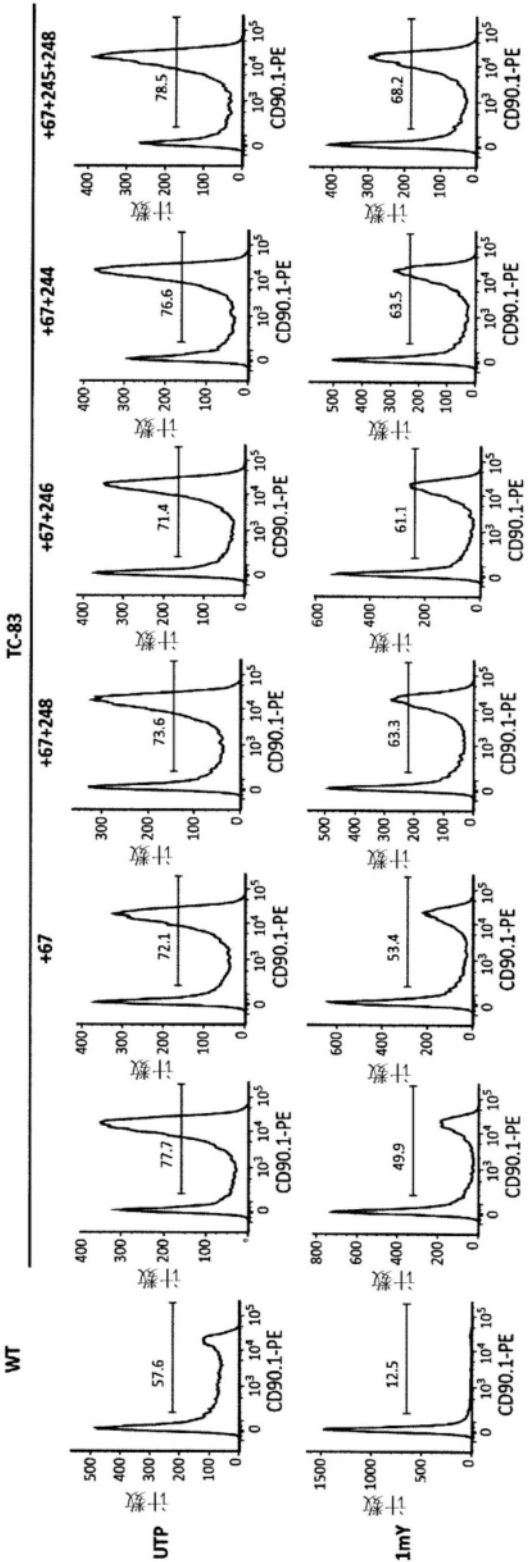


图8B

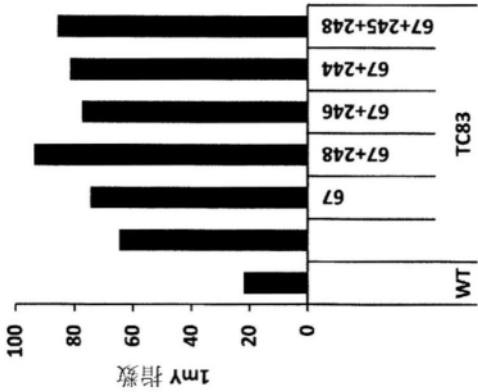


图8C

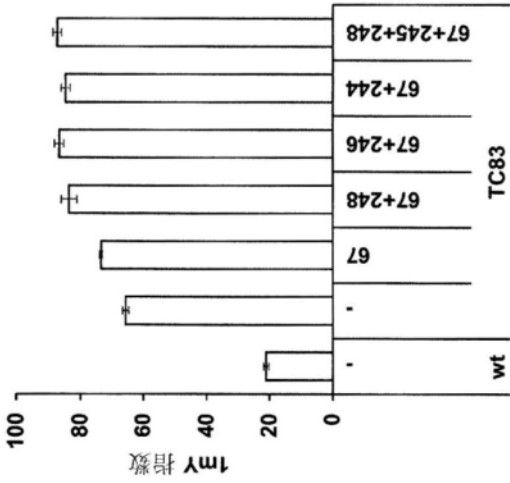


图8C-1

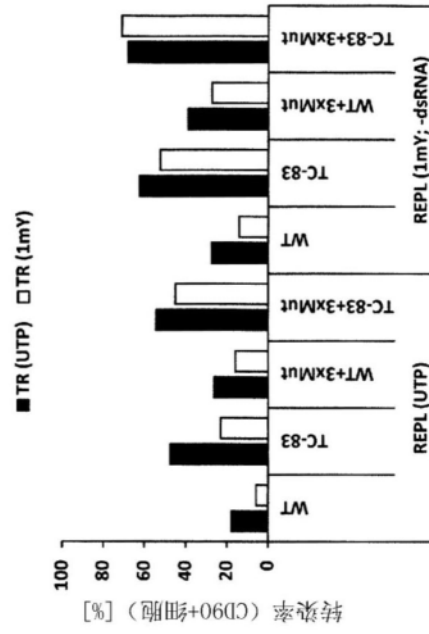


图9B

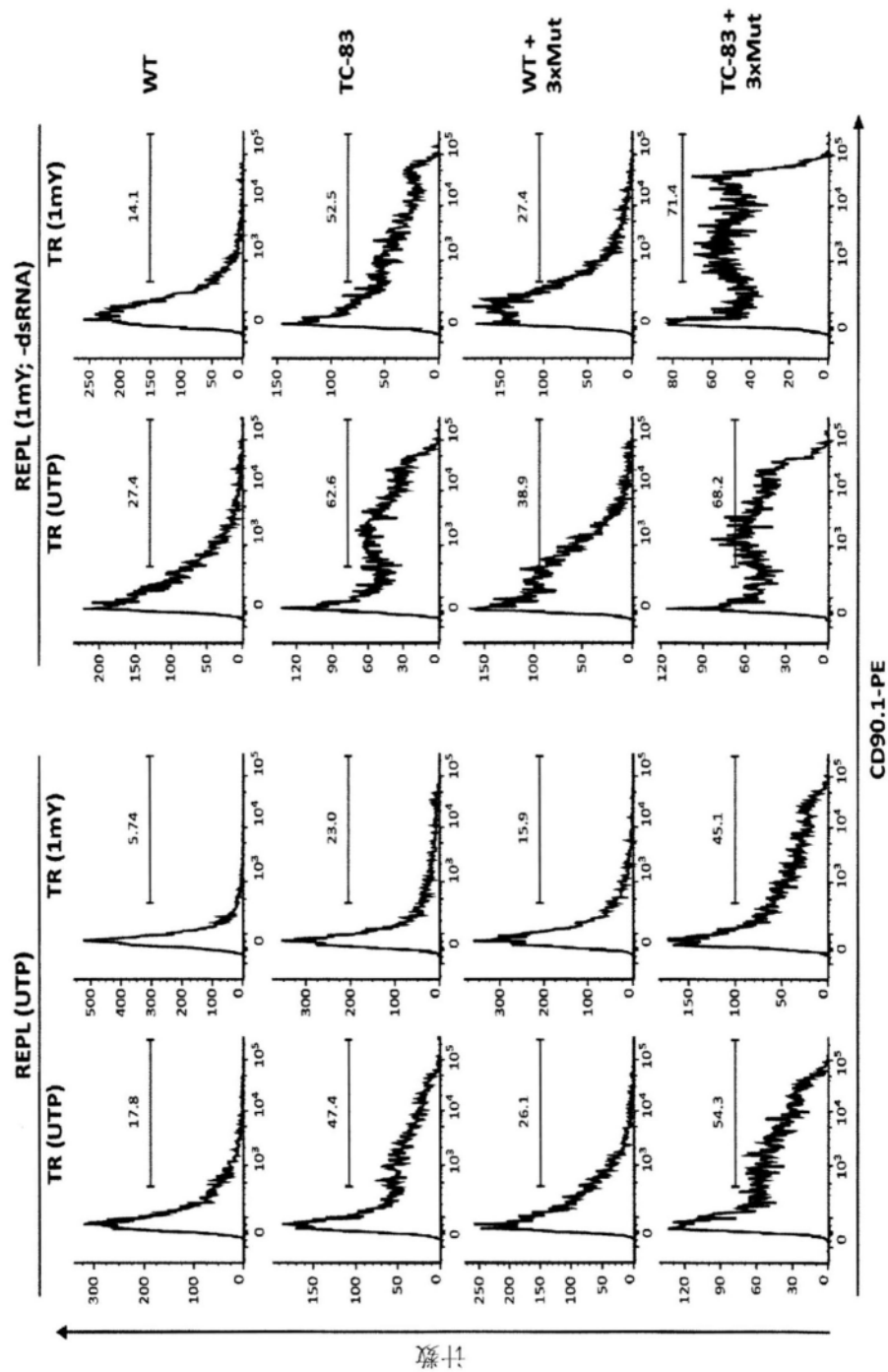


图9A

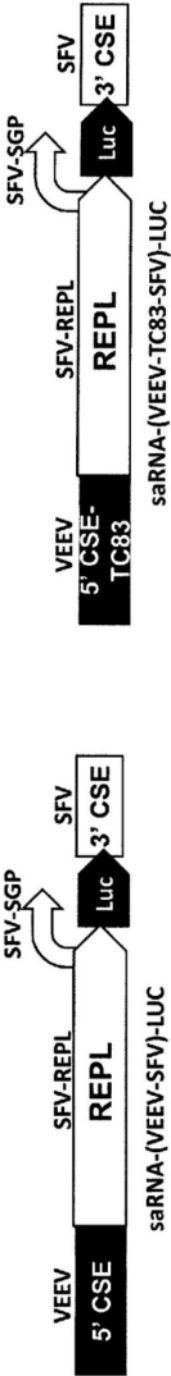


图10A

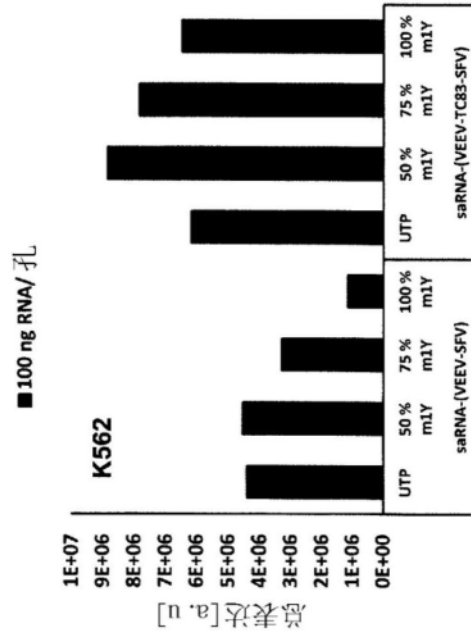


图10B

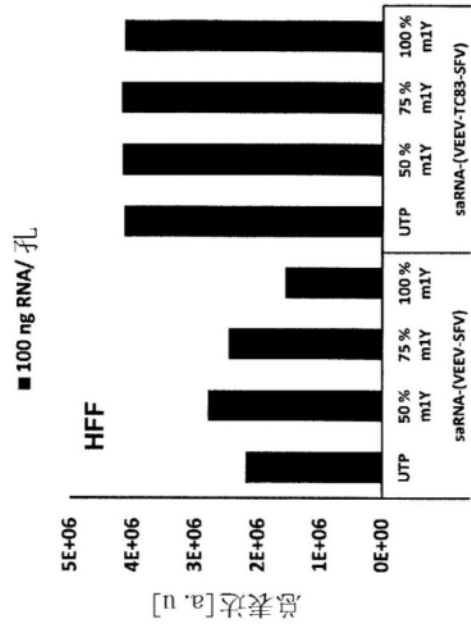


图10C

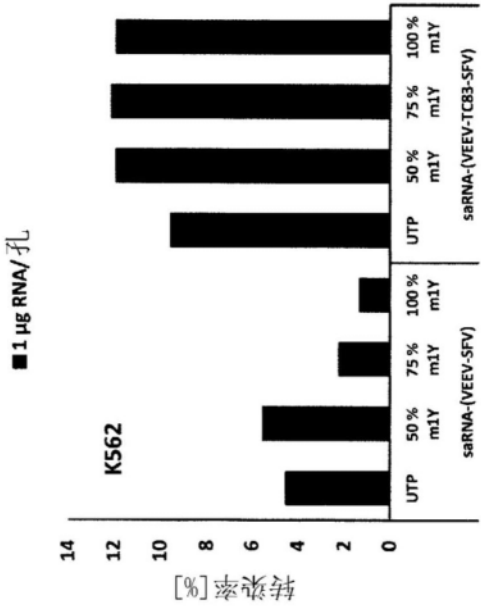


图10D

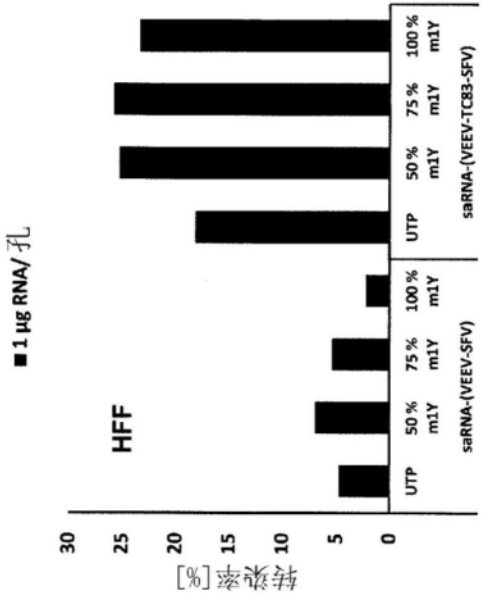


图10E

AUGGGCGGCG	CAUGAGAGAA	GCCCAGACCA	AUUACCUACC	CAAA <u>AUGGAG</u>	AAAGUUCACG	60
UUGACAUCGA	GGAAGACAGC	CCAUUCCUCA	GAGCUUUGCA	GCGGAGCUUC	CCGCAGUUUG	120
AGGUAGAAGC	CAAGCAGGUC	ACUGAUAAUG	ACCAUGC ^U AA	UGCCAGAGCG	UUUUCGCAUC	180
UGGCUUCAAA	ACUGAUCGAA	ACGGAGGUGG	ACCCAUCCGA	CACGAUCCUU	GACAUUGGAA	240
GUGCGCCCGC	CCGCAGAAUG	UAUUCUA				267

图11A

<u>GAUGGCGGC</u>	GCAUAAGAGA	AGCCCAGACC	AAUUAACCUAC	CCAAAUCGA	GAAAGUUCAC	60
GUUGAC <u>A</u> UCG	AGGAAGACAG	CCCAUUCUC	AGAGCUUUGC	AGCGGAGCUU	CCGGCAGUUU	120
GAGGUAGAAG	CCAAGGAGGU	GACUGAUAAU	CACCAUCCUA	UUGCCAGAGC	GUUUUCGCAU	180
CUGGCUUCAA	AACUGAUCGA	AACGGAGGUG	GACCCAUCCG	ACACGAUCCU	UGACAUUGGA	240
<u>AGUGCGCCCG</u>	CCCGCAGAAA	GUAUUCUACU	AGUGCCACCA	<u>UG</u>		282

图11B

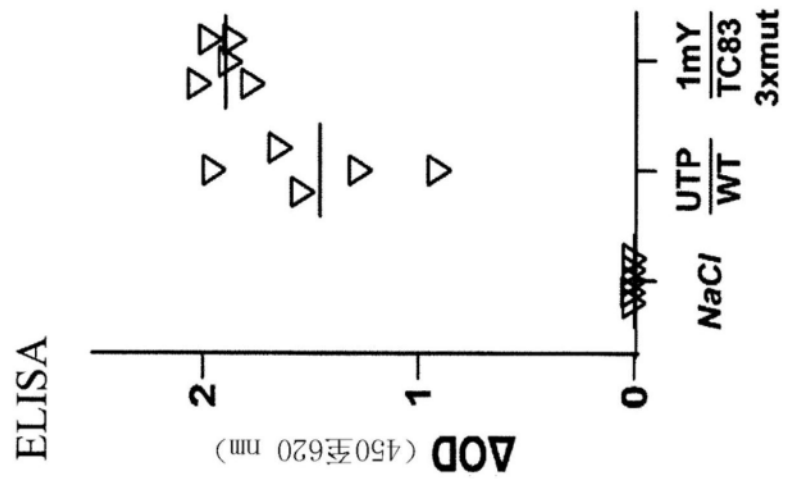


图12A

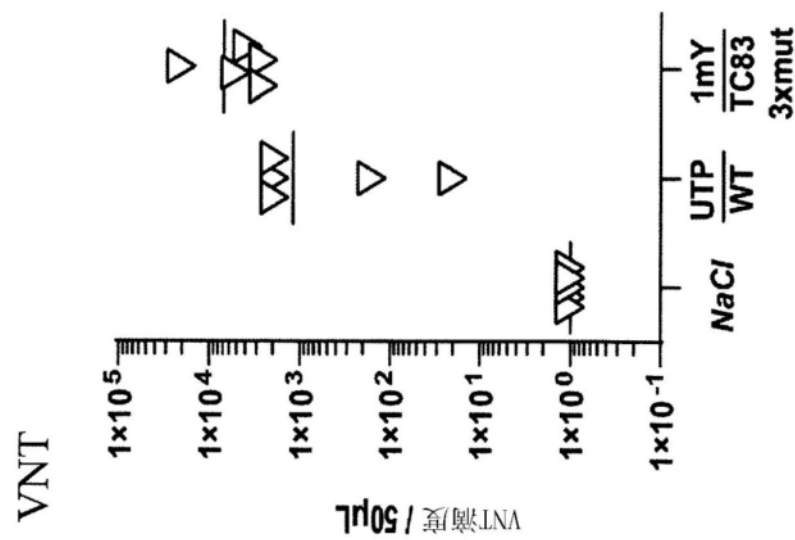


图12B

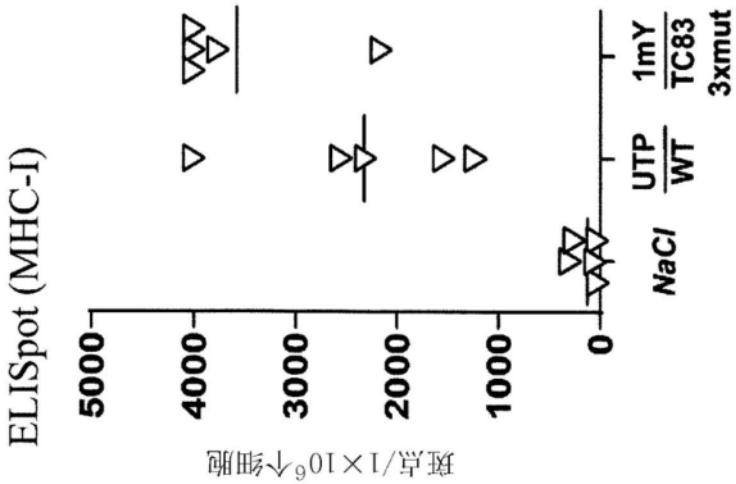


图12C

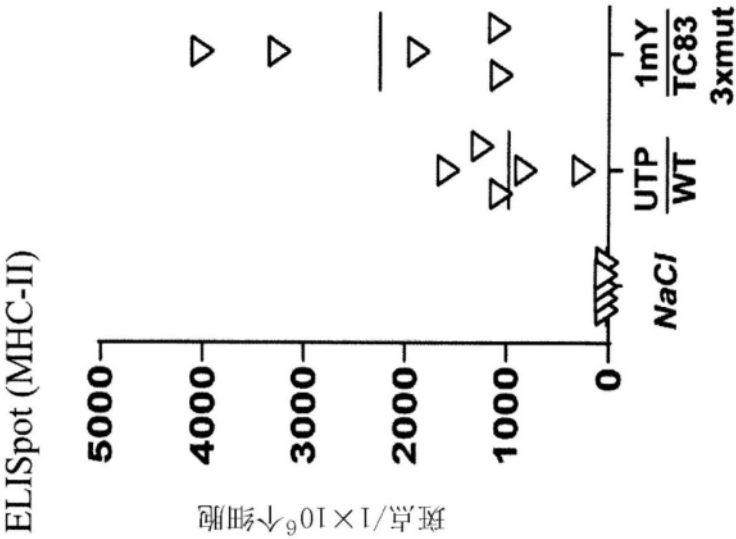


图12D

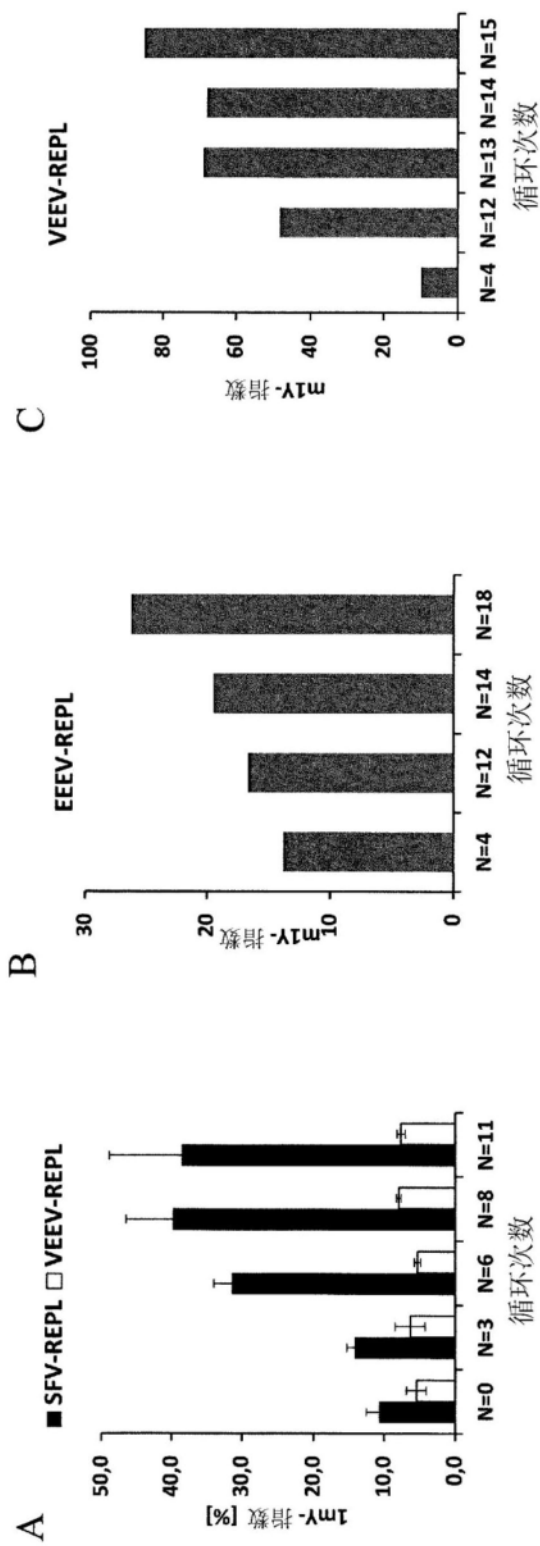


图13